Adaptive connectivity threshold of a locally regulated network using random measures on $[0,1]^2$-finite graphs

Khader Khadraoui and Ahmed Sid-Ali
Laval University, Department of Mathematics and Statistics
Québec, Qc, G1V 0A6, Canada

Abstract

We consider an interacting particle system for social networks associated with a finite and connected geometric graph. This system is a Markov process identified by $N_t \in \mathbb{N}$ independent particles with space states $x_t^1, \ldots, x_t^{N_t}$ in the unit square $[0,1]^2$ and with some interaction in continuous time $t \geq 0$. We study the behavior of isolated particles (nodes) distribution inside the network. Each node at any state x_t^i can be active or not in the network with a Bernoulli probability $p \in (0,1]$. Each pair being connected by affinity thanks to a step connection function if the interparticle distance $\|x_t^i - x_t^j\| \leq r_{\mathsf{af}}^\star$ for any $i \neq j$ and $r_{\mathsf{af}}^\star > 0$. We prove that when $r_{\mathsf{af}}^\star = \sqrt{\frac{(N_t)^{l-1}}{p\pi}}$ for $l \in (0,1)$, the number of isolated particles is governed by a Poisson distribution as the invitation rate of particles $\alpha \to \infty$ on large time interval. This offers a natural threshold for the construction of a $r_{\mathsf{af}}^\star$-neighborhood procedure tailored to the detection of isolated particles in a dynamic manner.

Keywords: Measure-valued processes; Interacting particle systems; Social networks; Isolated nodes; Geometric graphs; Monte Carlo.
MSC 2020 subject classifications: Primary 60D05, 05C80, 60J76, 60C05; secondary 60K35

1 Introduction

A problem common to many disciplines is that of adequately studying complex networks. Research on this problem occurs in applied mathematics (dynamic systems governed by ODEs and PDEs), probability and statistics (random graph models) and in computer science and engineering (statistical learning networks). A less familiar studying in this area is measure-valued processes in the class of interacting particle systems that are sometimes called also stochastic individual-based (or centered) models (IBMs). The theory of interacting particle systems is a quickly growing interdisciplinary area with a very broad spectrum of motivations and applications. In modern stochastic with diffusion or point processes and infinite-dimensional modeling, interacting particle models for the study of complex systems form a rich and powerful direction. Roughly, to present a briefly survey, interacting particle systems are widely used in nonlinear filtering problems (see, e.g., Carvalho et al. (1997); Crisan & Lyons (1997); Del Moral (1996)), in physics mainly for problems of fluid mechanics (see, e.g., McKean (1967); Sznitman (1984); Méléard & Roelly-Coppolett (1987)), in statistical mechanics (see, e.g., Liggett (1985); Spitzer (1969); Dobrushin (1971)), in recent years their application appears in ecology and biology (see,

e.g., Fournier & Méléard (2004); Finkelshtein et al. (2009); Tran (2008); Decreusefond et al. (2012)) and in the numerical approximations of Feynman-Kac measures (see, e.g., Del Moral & Doucet (2014).) For an account on the subject of interacting particle systems, we refer the reader to the recent monograph Dawson (2017) and references therein. We emphasize that the contact process on integer lattices or homogeneous trees which is one of the most studied interacting particle systems was introduced recently in the graph theory and probabilists started investigating this process on some families of random networks like configuration models, or preferential attachment graphs (see, e.g., Chatterjee & Durrett (2009); Mountford et al. (2013); Schapira & Valesin (2017).) In Durrett (2007) the reader can find a thorough introduction to some dynamic models on graphs.

In this work, we investigate a measure-valued process for social networks on a geometric random graph with adaptive connectivity threshold $r_{\mathsf{af}} > 0$ and evolving according to an interacting particle model as follows. Every particle not yet withdrawn is assumed to be capable to invite new particles with positive constant rate α. The withdrawal rate β is also assumed constant and present particles stay in the system during random exponential times with mean $1/\beta$ during which they interact with their neighbors following a local affinity function $\mathsf{aff}(x, y)$ of the distance between particles x and y. At the end of the presence period, the particle becomes removed and is no longer considered in the system. The whole system $\mathscr{N}$ represented by the set of all particles can attract and consequently recruits new particles by affinity with a state variable rate $w^{\mathsf{af}}(\cdot, \mathscr{N})$. Moreover, we assume in this paper that each particle can be active or not in the network with a Bernoulli probability $p \in (0, 1]$. With taking into account the activity of particles, we gain in realism and refine the results by including effects due to the striking difference of involvement between active and inactive particles and activity-dependent behaviors frequently observed in a wide range of real-world social networks. Accordingly, the size $N \in \mathbb{N}$ of the system is variable along the continuous time. Formally, we study a measure-valued process $(\mathscr{N}_t)_{t\geq 0}$ on a geometric graph G where its vertices V are characterized by a state x belonging to a closure space $\bar{\mathscr{D}}_G$ of an open connected subset $\mathscr{D}_G$ of $\mathbb{R}^d$, for some $d \geq 1$, and by their physical activity $a \in \{0, 1\}$. We set $\mathscr{X}_G := \bar{\mathscr{D}}_G \times \{0, 1\}$.

We denote by $\mathcal{M}_F(\mathscr{X}_G)$ the set of finite measures on $\mathscr{X}_G$. For $\mathscr{N} \in \mathcal{M}_F(\mathscr{X}_G)$ and for a real-valued bounded measurable function f, we define $\langle \mathscr{N}, f\rangle = \int_{\mathscr{X}_G} f d\mathscr{N}$. We denote also by $\mathcal{C}_b^{0,1}(\mathscr{X}_G, \mathbb{R})$ the space of continuous bounded real-valued functions on $\mathscr{X}_G$ with bounded continuous derivatives with respect to the activity variable a. Let recall that the set of cylindrical functions defined for each $\mu \in \mathcal{M}_F(\mathscr{X}_G)$ by $\Phi(\mu) = F\langle \mu, f\rangle$, with $F \in \mathcal{C}_b^1(\mathbb{R})$ and $f \in \mathcal{C}_b^{0,1}(\mathscr{X}_G)$ generates the set of bounded measurable functions on $\mathcal{M}_F(\mathscr{X}_G)$ as known from (Dawson, 1993, Theorem 3.2.6). Note that some additional notation is provided in sections which follow as needed. The measure-valued process $(\mathscr{N}_t)_{t\geq 0}$ on a geometric graph G is the Markov process on $\mathscr{X}_G$ and generator given, for any cylindrical function

Φ, by

$$\begin{aligned}\mathcal{L}\Phi(\mathscr{N}) = \langle \mathscr{N}, \partial_a f(\cdot)\rangle \Phi'(\mathscr{N}) &+ \beta \int_{\mathscr{X}_G} \Big\{\Phi(\mathscr{N} - \delta_{x,a}) - \Phi(\mathscr{N})\Big\}\mathscr{N}(dx, da) \\ &+ \alpha p \int_{\mathscr{X}_G} \mathscr{N}(dx, da) \int_{\mathbb{R}^d} \Big\{\Phi(\mathscr{N} + \delta_{x+z,1}) - \Phi(\mathscr{N})\Big\} k^{\mathsf{inv}}(x, z) dz \\ &+ \alpha(1-p) \int_{\mathscr{X}_G} \mathscr{N}(dx, da) \int_{\mathbb{R}^d} \Big\{\Phi(\mathscr{N} + \delta_{x+z,0}) - \Phi(\mathscr{N})\Big\} k^{\mathsf{inv}}(x, z) dz \\ &+ p \int_{\mathscr{X}_G} \Big\{ \int_{\bar{\mathscr{D}}_G} \big(\Phi(\mathscr{N} + \delta_{y,1}) - \Phi(\mathscr{N})\big) \mathsf{aff}(x, y) k^{\mathsf{af}}(y) dy \Big\} \mathscr{N}(dx, da) \\ &+ (1-p) \int_{\mathscr{X}_G} \Big\{ \int_{\bar{\mathscr{D}}_G} \big(\Phi(\mathscr{N} + \delta_{y,0}) - \Phi(\mathscr{N})\big) \mathsf{aff}(x, y) k^{\mathsf{af}}(y) dy \Big\} \mathscr{N}(dx, da),\end{aligned} \tag{1.1}$$

where $\delta_{x,a}$ denotes the Dirac measure at particle with state x and with activity a, $k^{\mathsf{inv}}(\cdot,\cdot)$ and $k^{\mathsf{af}}(\cdot)$ are two dispersion kernels needed for the recruitment of new particles. The above generator (1.1) prescribes that invited and recruited by affinity particles become members of the system with rate $\alpha + \int_{\mathscr{X}_G} \mathsf{aff}(x, y)\mathscr{N}(dx, da)$ and each particle disappears with rate β. Typically, the particles are related through a random network and will be represented by the vertices of an undirected geometric graph (Gilbert, 1961). We emphasize that between two neighbors we will place an edge if their distance is small than some threshold. The graph is nonoriented and an edge between x and y can be seen as two directed edges, one from x to y and the other from y to x. Vertices of the graph are interpreted as individuals in a social network. Note that representing networks by measure-valued equations and point processes has created rich advances in stochastic analysis of networks (see, e.g., Kang & Ramanan (2010); Perry & Wolfe (2013); Veitch & Roy (2015); Borgs et al. (2018); Caron & Fox (2017); Yuan et al. (2019); Atar et al. (2021)).

An important concept in the theory of complex networks is the connection between vertices which appears as a main issue in any connectivity study (Penrose, 2003). The purpose of the present paper is to analyse wisely the situation when the size of particles is almost surely large enough by taking the limit $\alpha \to \infty$ on large time interval. Particularly, we consider here some linear deterministic connectivity rule that consists in two members of the random system interact if and only if their distance does not exceed some threshold. More drastically, it is known that the main obstacle to connectivity inside certain networks is the existence of isolated members. In spite of not directly giving the optimal connectivity threshold in general case or at least in a larger class of connection functions aff, we prove herein that, when $d = 2$, the number of isolated particles has asymptotic Poisson distribution by employing the theory of concentration inequalities, a tailored discretization method of the system state space $\bar{\mathscr{D}}_G$ and some arguments from geometry and combinatorics. For one thing, we establish that the optimal threshold associated with this Poisson approximation is given by $r^{\star}_{\mathsf{af}} = \big((N_t)^{l-1}/p\pi\big)^{1/2}$ for $l \in (0, 1)$ and show that this result holds over connection functions that are zero beyond the threshold. This was previously known for random geometric graph and its variants when vertices are independently and uniformly distributed in the case $d = 2$ (Dette & Henze, 1989). Specifically, our result may be seen as a generalization of this Poisson approximation for the number of isolated vertices to random vertices generated from measure-valued processes or more generally interacting particle systems. Our generalization systematically encompasses the uniform case when $l \sim \log(\log(N_t))/\log(N_t)$ for N_t almost

surely large enough (when $\alpha \to \infty$ on large time interval.) We point out that the system of particles moving independently in $\bar{\mathscr{D}}_G$ according to a measure-valued process will be described by a rigorous pathwise representation using an inhomogeneous multivariate Poisson point measure with intensity measure $\lambda^\star$ of the form

$$\lambda^\star(dz, dy) = \alpha k^{\mathsf{inv}}(x, z)dz + \beta + \sum_{x \in V : x \sim y} \mathsf{aff}(x, y)k^{\mathsf{af}}(y)dy,\ \alpha > 0,\ \beta > 0, \tag{1.2}$$

for $x, y \in \bar{\mathscr{D}}_G$, $z \in \mathbb{R}^d$ and where $x \sim y$ means that x and y are neighbors. It is noteworthy to say that our strategy of proof is versatile in that it relies on quite useful methods and could easily be adapted to other contexts or other more general related measures. To our knowledge, the optimal threshold and the Poisson distribution that we establish are the first results in this direction and this progress creates as well other complex problems such as the extension to higher dimensions or other connection functions such as the Rayleigh fading function or functions that decay exponentially in some power of distance.

Another important consequence of the asymptotic distribution of isolated particles established in the present framework consists in the design of an adaptive dynamic method to detect isolated nodes. At first glance, this conceivable strategy enables us to detect groups and isolated members at any time $t \geq 0$ of course when our mild assumptions are satisfied. Notice that an important question investigated recently in graph theory consists in detecting the communities inside networks (see, e.g., Zhao et al. (2012); Arias-Castro & Verzelen (2014); Jin (2015); Bickel et al. (2015)). For instance, in the Erdös-Rényi graph context the stochastic block model (SBM) introduced in Holland et al. (1983) is usually used to model communities using the principle that the probability of a connection occurring between two members depends solely on their community membership. There are many extensions of the SBM for various applications, including the biological, communication and social networks (see, e.g., Bickel & Chen (2009); Snijders & Nowicki (1997); Park & Bader (2012)).

The outline of this paper can be delineated as follows: In Section 2, we present our preliminary results and the tools used throughout the paper. Particularly, we start by describing in detail the generative model and giving the exact Monte Carlo scheme for the computation of the interacting particle system. In Section 3, we state our main results. Most notably, we describe the discretization strategy used to study the distribution of vertices and the regularity of the network. Section 4 is devoted to some numerical evidence that involves the optimal threshold established in this paper in order to detect isolated vertices in continuous time. Section 5 contains the proofs of the main results of the paper where we establish the asymptotic adaptive threshold needed for the Poisson approximation law of isolated vertices behavior. We then discuss our results and the outlook in Section 6.

2 Preliminary results and tools

A geometric graph $G = (\mathscr{N}, r_{\mathsf{af}})$ will be understood as a set V of vertices and a set $E \subseteq \{\{x, y\} \subseteq V : x \neq y\}$ of edges generated from an interacting particle system $\mathscr{N}$ (a measure-valued process here). The geometric graphs we consider will always be connected where we connect two vertices by undirected link if and only if their distance is smaller than a certain neighborhood threshold of affinity $r_{\mathsf{af}} > 0$. For convenience, we will not explicitly treat graphs with loops (edges that start and end at the same vertex) and parallel edges between vertices, though one can define the measure-valued process on those

graphs as well and our results could then be readily adapted. We denote by $N = |G|$ the number of vertices of G; for a set A, we denote by $|A|$ the number of elements of A. We will often abuse notation by omitting the graph from some notations; for example, we may write the size N in place of N^G and the process on graph $\mathscr{N}$ in place of $\mathscr{N}^G$. The process $\mathscr{N}$ has a vector of unknown parameters, assumed to lie in $\theta = (p, \alpha, \beta, A_{\mathsf{f}}, r_{\mathsf{af}}, \theta^{\alpha}, \theta^{\mathsf{af}}) \in (0,1] \times (\mathbb{R}_+^*)^4 \times \mathbb{R}^{|\theta^{\alpha}|+|\theta^{\mathsf{af}}|}$ where θ^{α} denotes the parameter set of the invitation dispersion kernel $k^{\mathsf{inv}}(\cdot,\cdot)$ and θ^{af} denotes the parameter set of the affinity dispersion kernel $k^{\mathsf{af}}(\cdot,\cdot)$. We also denote by $\mathbb{P}_\theta$ a probability measure under which the measure-valued process with parameters θ is defined on the graph G (which will be clear from the context, as will the initial configuration of the process); later to avoid overly burdensome notation we omit θ from the notation as well. We denote by $\mathbb{E}_\theta$, or simply $\mathbb{E}$, the associated expectation.

2.1 Measure-valued process

We start with the description of a discrete stochastic measure-valued process model of a social network structured by spatial state and activity of its members. The description of this interacting particle system in continuous time consists of particles that invite other particles to join the system, interact by affinity with neighbors and leave the system with rates that may depend on spatial state and on their activity. During these events, the designed dynamic takes spatial-activity-structure into account. We follow and generalize the work of Khadraoui & Sid-Ali (2022) to which we refer for more details and full proofs. Particle is characterized by their state with values in $\bar{\mathscr{D}}_G$ (a closed subset of $\mathbb{R}^d$) and by their activity with values in $\{0,1\}$. We describe the whole system by a measure belonging to $\mathcal{M}_F(\mathscr{X}_G)$ (the set of finite measures on $\mathscr{X}_G$):

$$\mathscr{N}_t(dx, da) = \sum_{i=1}^{N_t} \delta_{x_t^i, a_t^i}(dx, da), \tag{2.1}$$

where each particle (vertex in G) is represented by a Dirac mass on its state and activity, and $N_t = \langle \mathscr{N}_t, 1\rangle = \int_{\mathscr{X}_G} \mathscr{N}_t(dx, da)$ is the system size at time t.

We shall describe the system by the evolution in time of the measure $\mathscr{N}_t$. For this, let define and introduce the following parameters and events:

(i) Each particle at state $x \in \bar{\mathscr{D}}_G$ gives invitation to a new particle for joining the system with rate $\alpha > 0$. With probability $p \in (0,1]$, the new particle is active with activity $a = 1$ and state $y = x + z$ where $z \in \mathbb{R}^d$ is drawn from a dispersion kernel $K^{\mathsf{inv}}(x, dz)$ assumed to satisfy

$$\int_{z\in\mathbb{R}^d, x+z\in\bar{\mathscr{D}}_G} K^{\mathsf{inv}}(x, dz) = 1 \quad \text{and} \quad \int_{z\in\mathbb{R}^d, x+z\notin\bar{\mathscr{D}}_G} K^{\mathsf{inv}}(x, dz) = 0.$$

With probability $(1-p) \in [0,1)$, the new particle is inactive with activity $a = 0$ and state $y = x+z$ where z is drawn similarly as previous (so that $x + z \in \bar{\mathscr{D}}_G$).

(ii) Each particle at state $x \in \bar{\mathscr{D}}_G$ leaves the system with rate $\beta > 0$.

(iii) Let $\mathsf{aff} : \bar{\mathscr{D}}_G \times \bar{\mathscr{D}}_G \to [0, \infty)$ be a real-valued kernel describing the interaction by affinity and assessing the strength of affinity between two particles. The whole system $\mathscr{N}$ may attract and

recruits new particle for joining the system with rate $w^{\mathsf{af}}(\cdot, \mathscr{N}) \in [0, \infty)$ given by

$$w^{\mathsf{af}}(y, \mathscr{N}) = \sum_{x \in V : x \sim y} \mathsf{aff}(x, y) = \int_{\mathscr{X}_G} \mathsf{aff}(x, y) \mathscr{N}(dx, da),$$

which models the total interaction on a new particle (y, a). With probability p, the new particle is active with activity $a = 1$ and state y drawn from an affinity dispersion kernel $K^{\mathsf{af}}(dy)$ assumed to satisfy

$$\int_{y \in \bar{\mathscr{D}}_G} K^{\mathsf{af}}(dy) = 1 \quad \text{and} \quad \int_{y \notin \bar{\mathscr{D}}_G} K^{\mathsf{af}}(dy) = 0.$$

With probability $(1 - p)$, the new particle is inactive with activity $a = 0$ and state y drawn from $K^{\mathsf{af}}(dy)$.

For clarity purposes, notice that in social networks at large being present and being active is totally different. To be active, the key is to update a social state frequently and engage in a conversation with fans and followers. Particularly, concerning the affinity kernel and due to some technical reasons in the asymptotic connectivity investigation we take

$$\mathsf{aff}(x, y) = A_{\mathsf{f}} \mathbb{1}_{\left\{ \|x - y\| \leq r_{\mathsf{af}} \right\}}(x, y), \tag{2.2}$$

where $\| \cdot \|$ denotes the Euclidean norm, $\mathbb{1}_A$ denotes the indicator function of the set A, $A_{\mathsf{f}} > 0$ denotes the maximum that interaction by affinity between two particles may reach and $r_{\mathsf{af}} > 0$ denotes the radius (affinity threshold) of the zone of interaction by affinity around each particle. Moreover, we assume that the dispersion kernels induce densities with respect to the Lebesgue measure such that $K^{\mathsf{inv}}(x, dz) = k^{\mathsf{inv}}(x, z)dz$ and $K^{\mathsf{af}}(dy) = k^{\mathsf{af}}(y)dy$.

Let us mention that different forms for the invitation, affinity and withdrawal rates can be proposed for more complex scenarios. For instance, the period devoted to the presence in the network before leaving may be very different from one particle to another and in this case a rate of the form $\beta(x, a)$ that takes into account of the state and activity makes more sense than a simple constant $\beta > 0$. In the same spirit, we may consider general rates $\alpha(x, a)$ and $\mathsf{aff}((x, a), (x', a'))$ depending simultaneously on state and activity as well.

In the sequel, we make the following assumption.

Assumption 1 (*i*) *The basic previous mechanisms (i), (ii) and (iii) of recruitment (by individual invitation and by affinity) and withdrawal are assumed to be independent.*

(*ii*) *We assume that there exist some positive reals* $\gamma_1 > 0$ *and* $\gamma_2 > 0$ *and two probability densities* $\tilde{k}^{\mathsf{inv}}$ *on* $\mathbb{R}^d$ *and* $\tilde{k}^{\mathsf{af}}$ *on* $\bar{\mathscr{D}}_G$ *such that, for all* $x \in \bar{\mathscr{D}}_G$,

$$k^{\mathsf{inv}}(x, z) \leq \gamma_1 \tilde{k}^{\mathsf{inv}}(z) \qquad \text{and} \qquad k^{\mathsf{af}}(y) \leq \gamma_2 \tilde{k}^{\mathsf{af}}(y).$$

There exists also a constant A_{f} *such that, for all* $x, y \in \bar{\mathscr{D}}_G$, $N \in \mathbb{N}$ *and* $\mathscr{N} \in \mathcal{M}_F(\mathscr{X}_G)$,

$$\mathsf{aff}(x, y) \leq A_{\mathsf{f}} \qquad \text{and then} \qquad w^{\mathsf{af}}(y, \mathscr{N}) \leq A_{\mathsf{f}} N.$$

The Assumption 1-(ii) is needed chiefly to avoid explosion phenomenon where we suppose that the spatial dependence of the introduced kernels and rates is bounded in some sense. Under Assumption 1, the measure-valued process $\mathscr{N}$ can be obtained as the unique strong solution of a stochastic differential equation (SDE) driven by the following five Poisson point measures (PPMs) corresponding to the dynamics described above.

Formally, let $(\Omega, \mathcal{A}, \mathbb{P})$ be a sufficiently large probability space and let $\mathscr{N}_0 \in \mathcal{M}_F(\mathscr{X}_G)$ be a random variable (initial condition of the process). On the probability space, let consider the PPMs $Q^{\alpha,1}(d\tau, di, dz, d\rho)$ defined on $[0,\infty)\times\mathbb{N}^*\times\mathbb{R}^d\times[0,1]$, $Q^{\alpha,0}(d\tau, di, dz, d\rho)$ defined on $[0,\infty)\times\mathbb{N}^*\times\mathbb{R}^d\times[0,1]$, $Q^{\beta}(d\tau, di, d\rho)$ defined on $[0,\infty)\times\mathbb{N}^*\times[0,1]$, $Q^{\mathsf{af},1}(d\tau, di, dy, d\rho)$ defined on $[0,\infty)\times\mathbb{N}^*\times\bar{\mathscr{D}}_G\times[0,1]$ and $Q^{\mathsf{af},0}(d\tau, di, dy, d\rho)$ defined on $[0,\infty)\times\mathbb{N}^*\times\bar{\mathscr{D}}_G\times[0,1]$ with respectively the intensity measures:

$$\begin{aligned}
n^{\alpha,1}(d\tau, di, dz, d\rho) &= p\alpha\gamma_1\tilde{k}^{\mathsf{inv}}(z)d\tau didzd\rho\\
n^{\alpha,0}(d\tau, di, dz, d\rho) &= (1-p)\alpha\gamma_1\tilde{k}^{\mathsf{inv}}(z)d\tau didzd\rho\\
n^{\beta}(d\tau, di, d\rho) &= \beta d\tau did\rho\\
n^{\mathsf{af},1}(d\tau, di, dy, d\rho) &= pA_{\mathsf{f}}\gamma_2\tilde{k}^{\mathsf{af}}(y)d\tau didyd\rho\\
n^{\mathsf{af},0}(d\tau, di, dy, d\rho) &= (1-p)A_{\mathsf{f}}\gamma_2\tilde{k}^{\mathsf{af}}(y)d\tau didyd\rho,
\end{aligned} \tag{2.3}$$

where $d\tau, dz, dy$ and $d\rho$ are the Lebesgue measures on $[0,\infty), \mathbb{R}^d, \bar{\mathscr{D}}_G$ and $[0,1]$ and di is the counting measure on $\mathbb{N}^*$. Suppose that the above PPMs and $\mathscr{N}_0$ are mutually independent. We also consider the canonical filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by the random objects $Q^{\alpha,1}$, $Q^{\alpha,0}$, Q^{β}, $Q^{\mathsf{af},1}$, $Q^{\mathsf{af},0}$ and $\mathscr{N}_0$. We denote by $\mathscr{N}_{t-}$ the network process at time t before any possible event.

We introduce now explicitly a SDE driven by the above PPMs, for which existence and uniqueness of the solution are stated. The solution is a Markov process, with generator (1.1) that corresponds to the dynamics described previously.

Definition 1 *For $t\in[0,\infty)$, $\mathscr{N}_t$ is described by:*

$$\begin{aligned}
\mathscr{N}_t =& \mathscr{N}_0 + \int_0^t\int_{\mathbb{N}^*}\int_{\mathbb{R}^d}\int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}}\mathbb{1}_{\{\rho\leq(k^{\mathsf{inv}}(x^i_{\tau-},z))/(\gamma_1\tilde{k}^{\mathsf{inv}}(z))\}}\delta_{x^i_{\tau-}+z,1}Q^{\alpha,1}(d\tau, di, dz, d\rho)\\
&+\int_0^t\int_{\mathbb{N}^*}\int_{\mathbb{R}^d}\int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}}\mathbb{1}_{\{\rho\leq(k^{\mathsf{inv}}(x^i_{\tau-},z))/(\gamma_1\tilde{k}^{\mathsf{inv}}(z))\}}\delta_{x^i_{\tau-}+z,0}Q^{\alpha,0}(d\tau, di, dz, d\rho)\\
&-\int_0^t\int_{\mathbb{N}^*}\int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}}\mathbb{1}_{\{\rho\leq(\beta/\beta)\}}\delta_{x^i_{\tau-},a^i_{\tau-}}Q^{\beta}(d\tau, di, d\rho)\\
&+\int_0^t\int_{\mathbb{N}^*}\int_{\bar{\mathscr{D}}_G}\int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}}\mathbb{1}_{\{\rho\leq(\mathsf{aff}(x^i_{\tau-},y)k^{\mathsf{af}}(y))/(A_{\mathsf{f}}\gamma_2\tilde{k}^{\mathsf{af}}(y))\}}\delta_{y,1}Q^{\mathsf{af},1}(d\tau, di, dy, d\rho)\\
&+\int_0^t\int_{\mathbb{N}^*}\int_{\bar{\mathscr{D}}_G}\int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}}\mathbb{1}_{\{\rho\leq(\mathsf{aff}(x^i_{\tau-},y)k^{\mathsf{af}}(y))/(A_{\mathsf{f}}\gamma_2\tilde{k}^{\mathsf{af}}(y))\}}\delta_{y,0}Q^{\mathsf{af},0}(d\tau, di, dy, d\rho).
\end{aligned} \tag{2.4}$$

Let $T_0 = 0$. Assume that the size N_t of the system at time $t\geq 0$ is finite. Assumption 1 ensures that the global jump rate at time t is upper bounded by $(\alpha+\beta+A_{\mathsf{f}})N_{t-}$ which is finite. Consequently, it is possible to define the sequence of jump times $(T_k)_{k\in\mathbb{N}^*}$ of $\mathscr{N}$ almost surely. Note here that this sequence is obviously increasing.

Theorem 1 *Assume that Assumption 1 holds. Then, for all $T > 0$, the $\mathbb{N}$-valued process $N_t = \langle \mathscr{N}_t, 1\rangle$ satisfies*

$$\mathbb{E}(\sup_{t\in[0,T]} N_t) < \mathbb{E}(N_0)e^{(\gamma_1\alpha + A_{\mathsf{f}}\gamma_2)T} < \infty, \tag{2.5}$$

and we have

$$T_\infty := \lim_{k\to+\infty} T_k = +\infty \qquad \mathbb{P}-\textit{almost surely}, \tag{2.6}$$

where T_∞ is well defined and SDE (2.4) admits a unique pathwise solution $(\mathscr{N}_t)_{t\geq 0} \in \mathcal{M}_F(\mathscr{X}_G)$.

Proof We can almost surely build a solution for SDE (2.4) algorithmically (its exact scheme will be given in Subsection 2.2) on $[0, T_\infty)$ where T_∞ is the random time defined above. To show existence of a strong solution for the SDE, we shall prove first (2.5) and second (2.6). For $n \in \mathbb{N}^*$, let consider the stopping time

$$\tau_n := \inf\{t \geq 0,\ N_t \geq n\}. \tag{2.7}$$

From (2.4), we can write for any $t \geq 0$,

$$\begin{aligned}
\langle \mathscr{N}_{\min(t,\tau_n)}, 1\rangle =& N_0 - \int_0^{\min(t,\tau_n)} \int_{\mathbb{N}^*} \int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}} \mathbb{1}_{\{\rho\leq(\beta/\beta)\}} Q^\beta(d\tau, di, d\rho) \\
&+ \int_0^{\min(t,\tau_n)} \int_{\mathbb{N}^*} \int_{\mathbb{R}^d} \int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}} \mathbb{1}_{\{\rho\leq(k^{\mathsf{inv}}(x^i_{\tau-},z))/(\gamma_1\tilde{k}^{\mathsf{inv}}(z))\}} \\
&\qquad \Big(Q^{\alpha,1}(d\tau, di, dz, d\rho) + Q^{\alpha,0}(d\tau, di, dz, d\rho)\Big) \\
&+ \int_0^{\min(t,\tau_n)} \int_{\mathbb{N}^*} \int_{\bar{\mathscr{D}}_G} \int_0^1 \mathbb{1}_{\{i\leq N_{\tau-}\}} \mathbb{1}_{\{\rho\leq(\mathsf{aff}(x^i_{\tau-},y)k^{\mathsf{af}}(y))/(A_{\mathsf{f}}\gamma_2\tilde{k}^{\mathsf{af}}(y))\}} \\
&\qquad \Big(Q^{\mathsf{af},1}(d\tau, di, dy, d\rho) + Q^{\mathsf{af},0}(d\tau, di, dy, d\rho)\Big).
\end{aligned} \tag{2.8}$$

By neglecting the nonpositive term in the right-hand side of (2.8), taking the expectation and applying the Fubini's theorem (since the integrants are bounded by Assumption 1), we obtain

$$\begin{aligned}
\mathbb{E}\Big(\sup_{s\in[0,\min(t,\tau_n)]} N_s\Big) &\leq \mathbb{E}(N_0) + (\gamma_1\alpha + A_{\mathsf{f}}\gamma_2)\int_0^t \mathbb{E}\Big(\sup_{u\in[0,\min(s,\tau_n)]} N_u\Big) ds \\
&\leq \mathbb{E}(N_0)\mathrm{e}^{(\gamma_1\alpha + A_{\mathsf{f}}\gamma_2)t},
\end{aligned} \tag{2.9}$$

where the upper bound (2.9) follows immediately from Gronwall's lemma. Obviously the bound (2.9) do not depend on n and consequently we deduce at first that $\lim_{n\to+\infty}\tau_n = +\infty$ $\mathbb{P}$-almost surely. We can verify this briefly by an absurd reasoning; If there exist $C < +\infty$ and a set $S_C \subset \Omega$ with positive probability such that, for all $\omega \in S_C$, $\lim_{n\to+\infty}\tau_n(\omega) < C$ then

$$\text{For all } T > C, \qquad \mathbb{E}\Big(\sup_{t\in[0,\min(T,\tau_n)]} N_t\Big) \geq n\mathbb{P}(S_C),$$

which contradicts (2.9). Now, we get rid of the stopping time τ_n by applying Fatou's lemma

$$\mathbb{E}\Big[\sup_{t\in[0,T]} N_t\Big] = \mathbb{E}\Big[\liminf_{n\to\infty} \sup_{t\in[0,\min(T,\tau_n)]} N_t\Big]$$
$$\leq \liminf_{n\to\infty} \mathbb{E}\Big[\sup_{t\in[0,\min(T,\tau_n)]} N_t\Big] \leq \mathbb{E}(N_0)\mathrm{e}^{(\gamma_1\alpha + A_{\mathsf{f}}\gamma_2)T} < +\infty,$$

which proves (2.5). By a second absurd reasoning we shall deduce that $T_\infty = \infty$ almost surely; Let us assume that there exist $\overline{C} < +\infty$ and a set $S_{\overline{C}} \subset \Omega$ with positive probability such that, for all $\omega \in S_{\overline{C}}$, $T_\infty(\omega) < \overline{C}$. For all $\omega \in S_{\overline{C}}$, if $\lim_{k\to+\infty} N_{T_k}(\omega) = +\infty$ then we have necessary, for all $n \in \mathbb{N}^*$, $\tau_n(\omega) \leq \overline{C}$ which contradicts our previous result about the limit of τ_n. It remains only to show that $\lim_{k\to+\infty} N_{T_k}(\omega) = +\infty$ and we proceed by assuming that this do not hold and there exist $n' \in \mathbb{N}^*$ and a set $S \subset S_{\overline{C}}$ with positive probability such that, for all $\omega \in S$ and all $k \in \mathbb{N}$, $N_{T_k}(\omega) < n'$. Hence, for all $\omega \in S$ we can design the jump times $(T_k)_{k\in\mathbb{N}^*}$ as a subsequence of points from a Poisson point process with intensity $(\alpha+\beta+A_{\mathsf{f}})n'$ where its only closure value is $+\infty$ $\mathbb{P}$-almost surely, which is absurd since we assume that for all $\omega \in S$, $T_\infty(\omega) < \overline{C}$. Finally, we conclude that (2.6) holds.

Let us now investigate the uniqueness of the SDE strong solution. For $k = 0$, the measure $\mathscr{N}_0$ is determined in unique manner by $(\mathscr{N}_0, Q) := (\mathscr{N}_0, Q^{\alpha,1}, Q^{\alpha,0}, Q^{\beta}, Q^{\mathsf{af},1}, Q^{\mathsf{af},0})$. We shall prove by recurrence that the sequence $(T_k, \mathscr{N}_{T_k})_{k\in\mathbb{N}}$ is determined in unique manner by $(\mathscr{N}_0, Q)$. For this, we assume that, for $k \in \mathbb{N}$, $(T_k, \mathscr{N}_{T_k})$ is determined in unique manner by $(\mathscr{N}_0, Q)$ and we shall prove that this holds also for $(T_{k+1}, \mathscr{N}_{T_{k+1}})$. Let us define:

$$\text{For all } t \geq T_k, \qquad \tilde{\mathscr{N}}^k_t(dx, da) = \sum_{i=1}^{N_{T_k}} \delta_{x^i_{T_k}, a^i_{T_k}}(dx, da),$$

which is determined in unique manner by $(T_k, \mathscr{N}_{T_k})$ or also by assumption by $(\mathscr{N}_0, Q)$. As a matter of fact, it turns out that

$$T_{k+1} = \inf\Big\{t > T_k,\ \Big(\int_{T_k}^t \int_{\mathbb{N}^*} \int_{\mathbb{R}^d} \int_0^1 \mathbb{1}_{\Big\{i\leq N_{T_k}, \rho \leq \frac{k^{\mathsf{inv}}(x^i_{\tau-}, z)}{\gamma_1 \bar{k}^{\mathsf{inv}}(z)}\Big\}} \Big[Q^{\alpha,1}(d\tau, di, dz, d\rho)$$
$$+ Q^{\alpha,0}(d\tau, di, dz, d\rho)\Big] + \int_{T_k}^t \int_{\mathbb{N}^*} \int_0^1 \mathbb{1}_{\Big\{i\leq N_{T_k}, \rho \leq (\beta/\beta)\Big\}} Q^{\beta}(d\tau, di, d\rho)$$
$$+ \int_{T_k}^t \int_{\mathbb{N}^*} \int_{\bar{\mathscr{D}}_G} \int_0^1 \mathbb{1}_{\Big\{i\leq N_{T_k}, \rho \leq \frac{\mathsf{aff}(x^i_{\tau-}, y) k^{\mathsf{af}}(y)}{A_{\mathsf{f}}\gamma_2 \bar{k}^{\mathsf{af}}(y)}\Big\}} \Big[Q^{\mathsf{af},1}(d\tau, di, dy, d\rho) + Q^{\mathsf{af},0}(d\tau, di, dy, d\rho)\Big]\Big) > 0\Big\},$$

where the jump time T_{k+1} is determined in unique manner by $(\mathscr{N}_0, Q)$. Since the measure-valued process is atomic, let consider now the atoms $(T_{k+1}, I_{k+1}, Z_{k+1}, \rho_{k+1})$, $(T_{k+1}, I_{k+1}, \rho_{k+1})$, and $(T_{k+1}, I_{k+1}, Y_{k+1}, \rho_{k+1})$ of respectively the PPMs $Q^{\alpha,1}$ with $Q^{\alpha,0}$, Q^{β}, and $Q^{\mathsf{af},1}$ with $Q^{\mathsf{af},0}$. For $\varphi : [0,1] \mapsto [0,\infty)$, it is worth

pointing out that

$$\begin{aligned}\mathscr{N}_{T_{k+1}}(dx,da) =& \tilde{\mathscr{N}}^{k}_{T_{k+1}}(dx,da) + \mathbb{1}_{\left\{\varphi(\rho_{k+1})\leq \alpha\times p\times k^{\mathsf{inv}}\left(x^{I_{k+1}}_{T_{k+1}-},Z_{k+1}\right)\right\}}\delta_{x^{I_{k+1}}_{T_{k+1}-}+Z_{k+1},1}(dx,da)\\ &+\mathbb{1}_{\left\{\varphi(\rho_{k+1})\leq \alpha\times(1-p)\times k^{\mathsf{inv}}\left(x^{I_{k+1}}_{T_{k+1}-},Z_{k+1}\right)\right\}}\delta_{x^{I_{k+1}}_{T_{k+1}-}+Z_{k+1},0}(dx,da)\\ &-\mathbb{1}_{\{\varphi(\rho_{k+1})\leq\beta\}}\delta_{x^{I_{k+1}}_{T_{k+1}-},a^{I_{k+1}}_{T_{k+1}-}}(dx,da)\\ &+\mathbb{1}_{\left\{\varphi(\rho_{k+1})\leq p\times \mathsf{aff}\left(x^{I_{k+1}}_{T_{k+1}-},Y_{k+1}\right)\times k^{\mathsf{af}}(Y_{k+1})\right\}}\delta_{Y_{k+1},1}(dx,da)\\ &+\mathbb{1}_{\left\{\varphi(\rho_{k+1})\leq (1-p)\times \mathsf{aff}\left(x^{I_{k+1}}_{T_{k+1}-},Y_{k+1}\right)\times k^{\mathsf{af}}(Y_{k+1})\right\}}\delta_{Y_{k+1},0}(dx,da),\end{aligned}$$

depends in unique manner on $(T_{k+1}, I_{k+1}, Z_{k+1}, Y_{k+1}, \rho_{k+1})$ which means that $\mathscr{N}_{T_{k+1}}$ is determined in unique manner by $(\mathscr{N}_0, Q)$. Thus, uniqueness holds since we have no choice in the construction which completes the proof. □

Remark 1 *Note that if we define the extinction time as the stopping time:*

$$\tau_0^\star := \inf\{t \geq 0, N_t = 0\},$$

with the convention $\inf \varnothing = \infty$*, then before* $\tau_0^\star$ *the infinitesimal generator is given by (1.1). After the extinction time* $(\mathscr{N}_t)_{t\geq\tau_0^\star}$ *is the null measure, i.e., the system does not contain any particle and the infinitesimal generator is simply reduced to null measure.*

2.2 About simulation and graphical construction

This subsection is devoted to design an exact Monte Carlo algorithm that simulates the measure-valued process and provides its empirical representation. Let $T > 0$ and let us consider the evolution system on the compact time interval $[0, T]$.

At time $t = 0$, we have a (possibly random) distribution $\mathscr{N}_0 \in \mathcal{M}_F(\mathscr{X}_G)$. Each particle (located at some $x \in \bar{\mathscr{D}}_G$ and with some activity $a \in \{0, 1\}$) has five independent exponential clocks: an invitation of active particle clock with parameter αp, an invitation of inactive particle clock with parameter $\alpha(1-p)$, a withdrawal clock with parameter β, a recruitment by affinity of active particle clock with parameter $p\mathsf{aff}(x,\cdot)$ and a recruitment by affinity of inactive particle clock with parameter $(1-p)\mathsf{aff}(x,\cdot)$. We emphasize that we compute the global clock of events that controls all point mechanisms by:

$$\mathsf{H}_t = \mathsf{h}_t^{\alpha,1} + \mathsf{h}_t^{\alpha,0} + \mathsf{h}_t^{\mathsf{af},1} + \mathsf{h}_t^{\mathsf{af},0} + \mathsf{h}_t^{\beta}, \tag{2.10}$$

where

$$\mathsf{h}_t^{\alpha,1} = \alpha p N_t, \quad \mathsf{h}_t^{\alpha,0} = \alpha(1-p)N_t, \quad \mathsf{h}_t^{\mathsf{af},1} = pA_{\mathsf{f}}N_t, \quad \mathsf{h}_t^{\mathsf{af},0} = (1-p)A_{\mathsf{f}}N_t, \text{ and } \quad \mathsf{h}_t^{\beta} = \beta N_t.$$

The pathwise definition (2.4) of the measure-valued process leads to the following simulation algorithm:

(*i*) To choose which event occurs, at each time step, we compute the probability of each event. Let T_{k-1} be the time of the last event and $\mathscr{N}_{T_{k-1}}$ be the corresponding random system. We simulate T_k and $\mathscr{N}_{T_k}$ by computing

$$T_k = T_{k-1} + S_k,$$

where $S_k \sim \mathrm{Exp}(\mathsf{H}_{T_{k-1}})$ and we set

$$\mathscr{N}_t = \mathscr{N}_{T_{k-1}}, \qquad \text{for } \ t \in [T_{k-1}, T_k).$$

(*ii*) We compute the probability of each event by:

$$\rho_k^{\alpha,1} = \frac{\mathsf{h}^{\alpha,1}_{T_{k-1}}}{\mathsf{H}_{T_{k-1}}}, \quad \rho_k^{\alpha,0} = \frac{\mathsf{h}^{\alpha,0}_{T_{k-1}}}{\mathsf{H}_{T_{k-1}}}, \quad \rho_k^{\mathsf{af},1} = \frac{\mathsf{h}^{\mathsf{af},1}_{T_{k-1}}}{\mathsf{H}_{T_{k-1}}}, \quad \rho_k^{\mathsf{af},0} = \frac{\mathsf{h}^{\mathsf{af},0}_{T_{k-1}}}{\mathsf{H}_{T_{k-1}}}, \qquad \text{and} \quad \rho_k^{\beta} = \frac{\mathsf{h}^{\beta}_{T_{k-1}}}{\mathsf{H}_{T_{k-1}}}.$$

(*iii*) We draw the candidate event as follows:

a. With probability $\rho_k^{\alpha,1}$ an invitation of active particle event occurs. We draw a present particle $(x^i_{T_{k-1}}, a^i_{T_{k-1}})$ where $i \sim U\{1, \dots, N_{T_{k-1}}\}$. We draw $z \in \mathbb{R}^d$ with the dispersal kernel $K^{\mathsf{inv}}(x^i_{T_{k-1}}, dz)$ and we add, with probability $\frac{k^{\mathsf{inv}}(x^i_{\tau_-}, z)}{\gamma_1 \tilde{k}^{\mathsf{inv}}(z)}$, a new particle to state $x^i_{T_{k-1}} + z$ (its activity is $a = 1$.) Finally, we set $\mathscr{N}_{T_k} = \mathscr{N}_{T_{k-1}} + \delta_{x^i_{T_{k-1}}+z,1}$.

b. With probability $\rho_k^{\alpha,0}$ an invitation of inactive particle event occurs and its description is similar to the previous event in (*iii*)-*a*.

c. With probability ρ_k^{β} a withdrawal event occurs. We draw a particle $(x^i_{T_{k-1}}, a^i_{T_{k-1}})$ where $i \sim U\{1, \dots, N_{T_{k-1}}\}$ and we set $\mathscr{N}_{T_k} = \mathscr{N}_{T_{k-1}} - \delta_{x^i_{T_{k-1}}, a^i_{T_{k-1}}}$.

d. With probability $\rho_k^{\mathsf{af},1}$ a recruitment by affinity of active particle event occurs. We draw an empty state y with the affinity kernel $K^{\mathsf{af}}(dy)$ and a particle $(x^i_{T_{k-1}}, a^i_{T_{k-1}})$ where $i \sim U\{1, \dots, N_{T_{k-1}}\}$. We reject the event with probability

$$1 - \frac{\mathsf{aff}(x^i_{T_{k-1}}, y) k^{\mathsf{af}}(y)}{A_{\mathsf{f}} \gamma_2 \tilde{k}^{\mathsf{af}}(y)},$$

otherwise we add a new particle at state y and we set $\mathscr{N}_{T_k} = \mathscr{N}_{T_{k-1}} + \delta_{y,1}$.

e. With probability $\rho_k^{\mathsf{af},0}$ a recruitment by affinity of inactive particle event occurs and its description is similar to the previous event in (*iii*)-*d*.

We introduce now the configuration with geometric graphs or threshold graphs and describe the dynamic of the interacting particle system $\mathscr{N}_{T_k}$, for $k \in \mathbb{N}$, on such graphs. Usually, graphs are represented as matrices with integer entries where to each graph corresponds an adjacency matrix, the (x, y)th coefficient of which is the number of edges between the vertices x and y. We emphasize that defining mathematically the distribution of a random graph amounts to choosing a sigma-field and a probability measure on the space $\mathbb{N}^{\mathbb{N}^* \times \mathbb{N}^*}$. Accordingly, we are interested in constructing random geometric graph $G_k = (\mathscr{N}_{T_k}, r_{\mathsf{af}})$, for $k \in \mathbb{N}$, that might be realistic in describing communities and

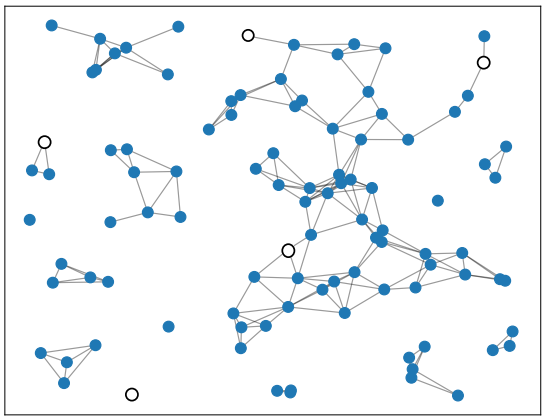


Figure 1: Geometric graph shown interacting particle system $\mathscr{N}_{T_k}$ at time T_k, for $k \in \mathbb{N}$, and in $\bar{\mathscr{D}}_G = [0,1]^2$, here $r_{\mathsf{af}} = 0.125$. Particles indicated by 100 small circles (vertices) connected by linear lines (edges) if the interparticle distance is small than 0.125. Active particles (here of size 95) are indicated by full circles and inactive particles (here of size 5) are indicated by empty circles. In total, there are 5 isolated particles (4 active and 1 inactive).

isolated vertices at each time T_k. By isolated we mean vertices that are not connected by edges with active vertices or at most connected only to inactive vertices (see Figure 1 for linkage procedure.) Formally, define $B(x, r_{\mathsf{af}}) = \{y \in \bar{\mathscr{D}}_G : \|x - y\| \leq r_{\mathsf{af}}\}$ to be generic notation for a ball with center x and radius $r_{\mathsf{af}} > 0$. We say that particle (x, a) is isolated if there are not active particles in the ball $B(x, r_{\mathsf{af}})$ except (x, a) itself.

The graph of relationships between particles is, in fact, relevant for studying the propagation of affinity and at large connectivity. The minimal information consists in the state $\mathscr{N}_{T_k}$ of the system from which we deduce the system size, the state-activity of vertices in $\bar{\mathscr{X}}_G$ and the set of edges for every active or inactive vertex. Thus, each vertex of determined types comes with a given number of edges.

3 Main results

Our chief motivation in this paper is to investigate the distribution of isolated vertices and to improve the assessment of the choice for the radius of affinity $(r_{\mathsf{af}}^k)_{k\geq 1}$ most particularly in function of the number of vertices $(N_{T_k})_{k\geq 1}$. For a formal construction, the study typically goes as follows: one increases the invitation rate α such that when $\alpha \to \infty$ and $k \to \infty$ then, for $\omega \in \Omega$, we have necessary

$$N_{T_k}(\omega) \to \infty, \qquad \mathbb{P}\text{-almost surely}, \tag{3.1}$$

fixes the other components of θ and some sequence of graphs $(G_k)_{k\geq 1}$ (usually converging or increasing, in some sense, to an infinite graph, or belonging to some class of random graphs), and then studies the asymptotic behavior of the random number of isolated vertices. With the generality that the results which follow are stated, we will restrict $\bar{\mathscr{D}}_G = [0,1]^2$ from now on due to technical convenience: it

allowed for the application of a certain splitting of unit square into small open boxes (see Lemma 1 below) and of geometric arguments for some decomposition of the domain $\bar{\mathscr{D}}_G$ (see Figure 6 below). To avoid technicalities arising from irregularities around the borders of $[0,1]^2$, we consider the unit square as a torus. Furthermore, it is well known that a fundamental question about any network is whether or not it is connected. We say herein that our social network is connected if each inactive vertex is adjacent to at least one active vertex together with all active vertices form a connected network.

Let $k_0 \geq 1$ be a large integer and consider the data $(\mathbf{x}_k, \mathbf{a}_k) = \big((x_{T_k}^1, a_{T_k}^1), \ldots, (x_{T_k}^{N_{T_k}}, a_{T_k}^{N_{T_k}})\big)$, for $1 \leq k \leq k_0$. Straightforwardly, we can observe during the time interval $[0, T_{k_0}]$ the sequences $(\mathbf{x}_k, \mathbf{a}_k)_{1\leq k\leq k_0}$ and $(N_{T_k})_{1\leq k\leq k_0}$ of random variables respectively in $[0,1]^{2\times N_{T_k}} \times \{0,1\}^{N_{T_k}}$ and $\mathbb{N}$, both adapted to filtration $(\bar{\mathcal{F}}_k)_{1\leq k\leq k_0}$. Our aim is to establish the optimal threshold of affinity that enables us to detect automatically isolated vertices in order to study their distribution at each time T_k based on the observation of $(\mathbf{x}_1, \mathbf{a}_1, \ldots, \mathbf{x}_{k_0}, \mathbf{a}_{k_0})$ and $(N_{T_1}, \ldots, N_{T_{k_0}})$, where $k_0 \geq 1$ is a finite $(\mathcal{F}_k)$-stopping time.

Remark 2 *The results given from now on are stated in a setting where one observes the network recorded data $(\mathbf{x}_k, \mathbf{a}_k, N_{T_k})_{1\leq k\leq k_0}$ with k_0 a stopping time. It is worth pointing out that this contains the usual case $k_0 \equiv 1$ and $N_{T_1} \equiv n$, where n is a fixed sample size. This strategy includes situations where the statistician decides to stop the recording process of network data according to some design of experiment rule.*

The analysis in the sequel is conducted under the following assumption.

Assumption 2 (A2) *There is a $(\mathcal{F}_k)$-adapted sequence of functions $\big(\mathfrak{f}_k(N_{T_k})\big)_{k\geq 1}$ of positive random variables. This sequence has a limit assumed to be known such that for $\mathfrak{f}_k : \mathbb{N} \mapsto \mathbb{R}$*

(*i*) *For all $k \geq 1$,*

$$\psi_k(N_{T_k}) := \log\Big(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\Big) + \mathfrak{f}_k(N_{T_k}) + \frac{\log(\alpha_k^\star(N_{T_k}))}{(N_{T_k}-1)^{-1}} - \alpha_k^\star(N_{T_k}) > 0 \quad \mathbb{P}-a.s., \tag{3.2}$$

where the sequence of functions $(\alpha_k^\star)_{k\geq 1}$ will be stated later on in (3.8) and depends on, among others, the parameters $\alpha, \beta, A_{\mathsf{f}}, r_{\mathsf{af}}, k^{\mathsf{inv}}, k^{\mathsf{af}}$.

(*ii*) *For all $k \geq 1$, one has $\mathbb{P}$-almost surely*

$$\lim_{\alpha,k\to\infty} \frac{\psi_k(N_{T_k})}{N_{T_k}} = 0, \qquad \textit{and} \qquad \lim_{\alpha,k\to\infty} \psi_k(N_{T_k}) = +\infty. \tag{3.3}$$

Let us make some comments on Assumption A2 now. The first condition (3.2) allows us to overcome the problem of the optimal affinity threshold assessment in the asymptotic study which follows whereas the second condition (3.3) means that ψ_k (a function that depends on the current size of the network) does not grow to ∞ more faster than N_{T_k} when both α and k tend to ∞.

3.1 Discretization and regularity

We will now state a few results concerning an in-depth analysis of our state space discretization and our key assumption about the asymptotic regularity of vertices states in the unit square. These results will be the basic tools in our proofs. The first two results are new but elementary and for the sake of completeness we sketch their proof, as they are short.

Define $B(x) := B(x, a_{\mathsf{f}}) \cap \bar{\mathscr{D}}_G$ and let denote by $\mathrm{vol}(x)$ the volume of $B(x)$ or in other words the Lebesgue measure of the measurable set $B(x)$. Straightforwardly, we have, for all $x \in \bar{\mathscr{D}}_G$, $\mathrm{vol}(x) \leq \pi r_{\mathsf{af}}^2$. For a purely notational reason, let E_{ik} be the event that $x^i_{T_k}$ is isolated and F_{ik} be the event that $x^i_{T_k}$ is an active isolated vertex for $i \in \{1, \ldots, N_{T_k}\}$ and $k \in \{1, \ldots, k_0\}$.

The proof relies heavily on some levels of discretization of the unit square into smaller subsquares (open boxes). Assume that $\bar{\mathscr{D}}_G = [0,1]^2$ is compact in $\mathbb{R}^2$ and consider the partitioning of $\bar{\mathscr{D}}_G$ into a family $\{\boldsymbol{C}_\ell\}_{\ell=1,\ldots,L_k}$ of disjoint boxes of $\mathbb{R}^2$ with side $r/\sqrt{2}$ that we need to cover the state space $\bar{\mathscr{D}}_G$ where $r \leq r_{\mathsf{af}}$. In the course of the proofs we will condition on the locations of vertices $x^1_{T_k}, \ldots, x^{N_{T_k}}_{T_k}$ and assume that they are sufficiently regularly distributed. The probability that this holds is proved in Lemma 1 that relies on concentration bound of large deviations for Poisson random variables (see Boucheron et al. (2013) for more details). For a box $\boldsymbol{C}_\ell$, we have

$$\mathbb{E}\Big[|\boldsymbol{C}_\ell \cap \bar{\mathscr{D}}_G|\Big] = \frac{N_{T_k}}{L_k} = \frac{N_{T_k} \times r^2}{2}.$$

Here, we need a definition for the regularity of boxes $\{\boldsymbol{C}_\ell\}_{\ell=1,\ldots,L_k}$.

Definition 2 *Fix $\nu \in (0,1)$. A box $\boldsymbol{C}_\ell$ is called* ν-regular *at time T_k if one has*

$$\frac{(1-\nu)N_{T_k} r^2}{2} \leq |\boldsymbol{C}_\ell \cap \bar{\mathscr{D}}_G| \leq \frac{(1+\nu)N_{T_k} r^2}{2}.$$

The following lemma is useful to assess the probability that all boxes $\boldsymbol{C}_\ell$ are ν-regular for N_{T_k} almost surely large enough.

Lemma 1 *For $k \geq 1$, there exist $\delta'_k > 0$ and $\gamma_k > 0$ such that for all $\delta_k \in [0, \delta'_k]$ and if $r(N_{T_k}) = \gamma_k \sqrt{\log(N_{T_k})/N_{T_k}}$, then by taking $\alpha \to \infty$ and $k \to \infty$,*

$$\inf_{\ell=1,\ldots,L_k} \mathbb{P}(\boldsymbol{C}_\ell \text{ is } \nu - \text{regular}) \geq 1 - 2L_k N_{T_k}^{-\gamma_k^2 \delta_k^2/6}.$$

In particular, if $\gamma_k^2 > 6/\delta_k^2$ then

$$\mathbb{P}(\text{every box } \boldsymbol{C}_\ell \text{ is } \nu - \text{regular}) \xrightarrow{\alpha, k \to \infty} 1.$$

Proof Let $\mathcal{P}(\lambda_0)$ be a Poisson random distribution with mean λ_0. By a standard concentration bound for Poisson random variables, we have for $\delta_0 > 0$ and $\delta \in [0, \delta_0]$

$$\mathbb{P}(|\mathcal{P}(\lambda_0) - \lambda_0| \geq \delta \lambda_0) \leq 2\mathrm{e}^{-\lambda_0 \delta^2/3}. \tag{3.4}$$

For any box $\boldsymbol{C}_\ell$, the number of vertices $|\boldsymbol{C}_\ell \cap \bar{\mathscr{D}}_G|$ is distributed like a Poisson random variable with mean $N_{T_k} \times r^2/2$. By applying (3.4), we establish for $\delta_k' > 0$ and $\delta_k \in [0, \delta_k']$ that

$$\mathbb{P}\Big(\Big|\mathcal{P}\Big(\frac{N_{T_k} r^2}{2}\Big) - \frac{N_{T_k} r^2}{2}\Big| \geq \delta_k \frac{N_{T_k} r^2}{2}\Big) \leq 2\mathrm{e}^{-N_{T_k} r^2 \delta_k^2/2.3}. \tag{3.5}$$

Now, for every box and when $\alpha, k \to \infty$,

$$\begin{aligned}\mathbb{P}(\boldsymbol{C}_\ell \text{ is not } \nu - \text{regular}) &\leq 2\mathrm{e}^{-N_{T_k} r^2 \delta_k^2/6} \\ &\leq 2N_{T_k}^{-\gamma_k^2 \delta_k^2/6},\end{aligned}$$

since $r(N_{T_k}) = \gamma_k \sqrt{\log(N_{T_k})/N_{T_k}}$ for any $k \geq 1$. In addition, if there exists one box that is not ν-regular, then one of the L_k boxes has a number of vertices that is out of range, so that as $\alpha, k \to \infty$,

$$\begin{aligned}\mathbb{P}(\exists \boldsymbol{C}_\ell : \boldsymbol{C}_\ell \text{ is not } \nu - \text{regular}) &\leq 2 \times L_k \times N_{T_k}^{-\gamma_k^2 \delta_k^2/6} = 2 \times \frac{2N_{T_k}}{\gamma_k^2 \log(N_{T_k})} \times N_{T_k}^{-\gamma_k^2 \delta_k^2/6} \\ &\leq N_{T_k}^{1-\gamma_k^2 \delta_k^2/6+o(1)},\end{aligned}$$

which tends to zero provided that $\gamma_k^2 > 6/\delta_k^2$. □

For simplicity and from now on we make the following technical assumption.

Assumption 3 (A3) *Assume that, for all $k \geq 1$, $r_{\mathsf{af}}(N_{T_k}) \geq r(N_{T_k})$ $\mathbb{P}$-almost surely where $r(N_{T_k}) \sim \gamma_k \sqrt{\log(N_{T_k})/N_{T_k}}$ with $\gamma_k > \sqrt{6}/\delta_k$ and $\delta_k \in [0, \delta_k']$ and $\delta_k' > 0$. Assume also that there exists a $(\mathcal{F}_k)$-adapted sequence of multivariate functions $\big(g_k(x, N_{T_k})\big)_{k \geq 1}$, with $g_k : \bar{\mathscr{D}}_G \times \mathbb{N} \mapsto \mathbb{R}_+$, such that, for all $y, x_{T_k}^1, \ldots, x_{T_k}^{N_{T_k}} \in \bar{\mathscr{D}}_G$ we have*

$$\sum_{i=1}^{N_{T_k}} \mathbb{1}_{\{\|x_{T_k}^i - y\| \leq r_{\mathsf{af}}\}} = g_k(y, N_{T_k}) \sum_{i=1}^{N_{T_k}} \sum_{\ell=1}^{L_k} \mathbb{1}_{\{(\boldsymbol{C}_\ell \cap B(y)) \neq \varnothing\}} \mathbb{1}_{\boldsymbol{C}_\ell}(x_{T_k}^i),$$

with $g_k(y, N_{T_k}) \xrightarrow{\alpha, k \to \infty} 1$ $\mathbb{P}$- a.s. and where $\{\boldsymbol{C}_\ell\}_{\ell=1,\ldots,L_k}$ is a family of disjoint boxes with side $r(N_{T_k})/\sqrt{2}$ that covers $\bar{\mathscr{D}}_G$.

The Assumption A3 is mainly needed to assess the accuracy of the approximation of $|B(y)|$ (the number of vertices in the neighborhood of each vertex $(y, a) \in \bar{\mathscr{D}}_G \times \{0, 1\}$.) Notice that L_k is clearly finite (for all $k \geq 1$) as well as for each $x, y \in \boldsymbol{C}_\ell$, $\|x - y\| \leq r$. Notice also that the function g_k assesses the proportion of vertices inside the ball $B(y)$ from the whole number of vertices inside the boxes that intersect with $B(y)$. Such an assumption is realistic because, when N_{T_k} is almost surely large enough, the side of each box $r(N_{T_k})/\sqrt{2}$ is small enough which guarantees a fine mesh and thus $g_k \sim 1$ for all $y \in \bar{\mathscr{D}}_G$. More rigorously, we use the fact that all open of $\mathbb{R}^d$ is a countable reunion of open pavers. Furthermore, the assumption is realistic since asymptotically all the boxes contain vertices (proved in Lemma 1) and is needed in practice to discard regions with strong variations.

In order to use A3 for estimating the probability of isolated vertices, we need to make sure that the global affinity rate stays under control.

Lemma 2 *Admit A3. For any $x, y \in \bar{\mathscr{D}}_G$, $z \in \mathbb{R}^2$ and any bounded region $B(x)$, one consider the (finite) integral of intensity measure $\lambda_k^\star$ given by (1.2) over region $B(x)$*

$$\Lambda_k^\star(x) := \int_{B(x)} \lambda_k^\star = \int_{B(x)} \Big(\alpha k^{\mathsf{inv}}(x,z) + \beta + \sum_{x \in \mathscr{N}_{T_k}} \mathsf{aff}(x,y) k^{\mathsf{af}}(y)\Big) dx. \tag{3.6}$$

Then, there exists a constant $\tilde{c}$ in $B(x)$ such that for $\alpha, k \to \infty$

$$\Lambda_k^\star(x) = \Big(\alpha k^{\mathsf{inv}}(\tilde{c}, z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\Big) \mathrm{vol}(x), \tag{3.7}$$

where $\tilde{L}_k$ is an integer in finite non-empty set of integers $\mathscr{C}_k := \Big\{ \lfloor \frac{\pi(r_{\mathsf{af}}(N_{T_k}))^2}{2(r(N_{T_k}))^2} \rfloor, \ldots, \lfloor \frac{2\pi(r_{\mathsf{af}}(N_{T_k}))^2}{(r(N_{T_k}))^2} \rfloor \Big\}$.

Proof We apply A3 to show that direct computation of (3.6) gives (3.7). First, for any $y \in \bar{\mathscr{D}}_G$ we easily notice that the number of boxes that intersect with the bounded region $B(y)$ is between $\lfloor \frac{\pi(r_{\mathsf{af}}(N_{T_k}))^2/4}{(r(N_{T_k})/\sqrt{2})^2} \rfloor$ (when y is located at one corner of $\bar{\mathscr{D}}_G$) and $\lfloor \frac{\pi(r_{\mathsf{af}}(N_{T_k}))^2}{(r(N_{T_k})/\sqrt{2})^2} \rfloor$ (when $B(y)$ is fully contained in $\bar{\mathscr{D}}_G$.) For all $k \geq 1$, let $\tilde{L}_k$ denote the number of boxes that intersect with $B(y)$. When N_{T_k} is almost surely sufficiently large, every square that intersects with $B(y)$ is fully contained in $B(y)$ and the large number of boxes allows us to take advantage of the approximation $g_k(y, N_{T_k}) \sim 1$ for all $y \in \bar{\mathscr{D}}_G$. Now, since every square is ν-regular for every $y \in \bar{\mathscr{D}}_G$, it follows that

$$\begin{aligned}
\int_{B(x)} A_{\mathsf{f}} \sum_{i=1}^{N_{T_k}} \mathbb{1}_{\{\|x_{T_k}^i - y\| \leq r_{\mathsf{af}}\}} dx &= A_{\mathsf{f}} \int_{B(x)} \sum_{i=1}^{N_{T_k}} \sum_{\ell=1}^{L_k} \mathbb{1}_{\{(\boldsymbol{C}_\ell \cap B(y)) \neq \varnothing\}} \mathbb{1}_{\boldsymbol{C}_\ell}(x_{T_k}^i) dx \\
&= A_{\mathsf{f}} \int \sum_{i=1}^{N_{T_k}} \sum_{\ell=1}^{\tilde{L}_k} \mathbb{1}_{\{(\boldsymbol{C}_\ell \cap B(y)) \neq \varnothing\}} \mathbb{1}_{\boldsymbol{C}_\ell}(x_{T_k}^i) \mathbb{1}_{B(x)} dx \\
&= \Big(A_{\mathsf{f}} \tilde{L}_k \frac{N_{T_k} \times (r(N_{T_k}))^2}{2}\Big) \int \mathbb{1}_{B(x)} dx \\
&= \Big(A_{\mathsf{f}} \tilde{L}_k \frac{N_{T_k} \times \gamma_k^2 \log(N_{T_k})/N_{T_k}}{2}\Big) \mathrm{vol}(x),
\end{aligned}$$

and the claim follows easily by application of the mean-value theorem. □

We are now in position to compute the probability of an isolated vertex inside the network.

Proposition 1 *Under A3, for any $x_{T_k}^i, y \in \bar{\mathscr{D}}_G$ and any $z \in \mathbb{R}^2$, we have for $\alpha, k \to \infty$ and for $i, k \in \{1, \ldots, N_{T_k}\} \times \{1, \ldots, k_0\}$*

$$\begin{aligned}
\mathbb{P}(E_{ik}) =& \frac{\big(\alpha k^{\mathsf{inv}}(\tilde{c}, z) + \beta + \frac{1}{2} A_{\mathsf{f}} \tilde{L}_k \gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)\big)^{N_{T_k}-1}}{(N_{T_k}-1)!} \mathrm{e}^{-\left(\alpha k^{\mathsf{inv}}(\tilde{c}, z) + \beta + \frac{1}{2} A_{\mathsf{f}} \tilde{L}_k \gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)\right)} \\
&\times \int_{\bar{\mathscr{D}}_G} \Big(1 - p \mathrm{vol}(x_{T_k}^i)\Big)^{N_{T_k}-1} dx_{T_k}^i,
\end{aligned}$$

where $\tilde{L}_k \in \mathscr{C}_k$, $\alpha, \beta, A_{\mathsf{f}} > 0$, $0 < p \leq 1$, $k^{\mathsf{inv}}(\cdot, \cdot)$ and $k^{\mathsf{af}}(\cdot)$ are two probability densities on $\mathbb{R}^2$ and $\bar{\mathscr{D}}_G$.

Proof At first sight and for simplicity, let Ω^{ik} be the event that, for any $x^i_{T_k} \in \bar{\mathscr{D}}_G$ and for any $j \neq i$, $x^j_{T_k} \notin B(x^i_{T_k})$ or $x^j_{T_k} \in B(x^i_{T_k})$ but inactive (i.e., $a^j_{T_k} = 0$). By direct calculation we obtain

$$
\begin{aligned}
\mathbb{P}(E_{ik}) &= \int_{\bar{\mathscr{D}}_G} \mathbb{P}\Big(E_{ik}\Big| x^i_{T_k} \in \bar{\mathscr{D}}_G\Big) dx^i_{T_k} = \int_{\bar{\mathscr{D}}_G} \mathbb{P}(\Omega^{ik}) dx^i_{T_k} \\
&= \int_{\bar{\mathscr{D}}_G} \sum_{\ell=0}^{N_{T_k}-1} \Big((1-p)^\ell \frac{\big(\int_{B(x^i_{T_k})} \lambda^\star_k\big)^\ell}{\ell!} \mathrm{e}^{-\int_{B(x^i_{T_k})} \lambda^\star_k} \Big) \\
&\qquad \times \Big(\frac{\big(\int_{\bar{\mathscr{D}}_G \setminus B(x^i_{T_k})} \lambda^\star_k\big)^{N_{T_k}-1-\ell}}{(N_{T_k}-1-\ell)!} \mathrm{e}^{-\int_{\bar{\mathscr{D}}_G \setminus B(x^i_{T_k})} \lambda^\star_k} \Big) dx^i_{T_k} \\
&= \int_{\bar{\mathscr{D}}_G} \sum_{\ell=0}^{N_{T_k}-1} \Big[(1-p)^\ell \frac{\Big(\Big(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\Big) \mathrm{vol}(x^i_{T_k})\Big)^\ell}{\ell!} \\
&\qquad \times \mathrm{e}^{-\left(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\right) \mathrm{vol}(x^i_{T_k})} \\
&\qquad \times \frac{\Big(\Big(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\Big) \big(1 - \mathrm{vol}(x^i_{T_k})\big)\Big)^{(N_{T_k}-1-\ell)}}{(N_{T_k}-1-\ell)!} \\
&\qquad \times \mathrm{e}^{-\left(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\right) \left(1 - \mathrm{vol}(x^i_{T_k})\right)} \Big] dx^i_{T_k}.
\end{aligned}
$$

We left out the terms that cancel out which leads to

$$
\begin{aligned}
\mathbb{P}(E_{ik}) &= \frac{\mathrm{e}^{-\left(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\right)}}{(N_{T_k}-1)!} \int_{\bar{\mathscr{D}}_G} \sum_{\ell=0}^{N_{T_k}-1} \frac{(N_{T_k}-1)!}{\ell!(N_{T_k}-1-\ell)!} \\
&\qquad \times \Big((1-p)\big(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\big) \mathrm{vol}(x^i_{T_k}) \Big)^\ell \\
&\qquad \times \Big(\big(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\big) \big(1 - \mathrm{vol}(x^i_{T_k})\big) \Big)^{N_{T_k}-1-\ell} dx^i_{T_k} \\
&= \frac{\big(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\big)^{N_{T_k}-1}}{(N_{T_k}-1)!} \mathrm{e}^{-\left(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + A_{\mathsf{f}} \tilde{L}_k \frac{\gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)}{2}\right)} \\
&\qquad \times \int_{\bar{\mathscr{D}}_G} \Big(1 - p \mathrm{vol}(x^i_{T_k})\Big)^{N_{T_k}-1} dx^i_{T_k},
\end{aligned}
$$

where the last line is obtained from the binomial theorem. This completes the proof. □

It is worth pointing out that the result cited in Proposition 1 suggests that the probability of being a vertex isolated in the network is inversely proportional to the size of the network and at the same time to the volume of the ball portion (around the vertex) that intersects with $\bar{\mathscr{D}}_G$. Thus, this result seems intuitive. Now, we shall assess the probability that more than one vertex (say $\kappa \geq 2$ vertices) are isolated inside the network. For any $(x^1, \ldots, x^\kappa) \in (\bar{\mathscr{D}}_G)^\kappa$ and to shorten notation, we use from now on $B(x^1, \ldots, x^\kappa) := B(x^1) \cup \cdots \cup B(x^\kappa)$ and $\mathrm{vol}(x^1, \ldots, x^\kappa)$ for the volume of $B(x^1, \ldots, x^\kappa)$.

Proposition 2 *Under A3, for $\kappa \geq 2$ and $(x^1_{T_k}, \ldots, x^\kappa_{T_k}) \in (\bar{\mathscr{D}}_G)^\kappa$, we have for $\alpha, k \to \infty$ and $k \in \{1, \ldots, k_0\}$*

$$\mathbb{P}(E_{1k} \cap \cdots \cap E_{\kappa k}) \leq \frac{\left(\alpha k^{\mathsf{inv}}(\tilde{c}, z) + \beta + \frac{1}{2} A_{\mathsf{f}} \tilde{L}_k \gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)\right)^{N_{T_k} - \kappa}}{(N_{T_k} - \kappa)! \, \mathrm{e}^{\left(\alpha k^{\mathsf{inv}}(\tilde{c}, z) + \beta + \frac{1}{2} A_{\mathsf{f}} \tilde{L}_k \gamma_k^2 \log(N_{T_k}) k^{\mathsf{af}}(y)\right)}} \times \int_{(\bar{\mathscr{D}}_G)^\kappa} \Big(1 - p\mathrm{vol}(x^1_{T_k}, \ldots, x^\kappa_{T_k})\Big)^{N_{T_k} - \kappa} dx^1_{T_k} \cdots dx^\kappa_{T_k},$$

where $\tilde{L}_k \in \mathscr{C}_k$, $\alpha, \beta, A_{\mathsf{f}} > 0$, $0 < p \leq 1$, $k^{\mathsf{inv}}(\cdot, \cdot)$ and $k^{\mathsf{af}}(\cdot)$ are two probability densities on $\mathbb{R}^2$ and $\bar{\mathscr{D}}_G$.

Proof For any $\kappa \geq 2$ and $(x^1_{T_k}, \ldots, x^\kappa_{T_k}) \in (\bar{\mathscr{D}}_G)^\kappa$, let $\Omega_0^{\kappa k}$ be the event that $B(x^1_{T_k}, \ldots, x^\kappa_{T_k})$ does not contain any active vertex in states $x^{\kappa+1}_{T_k}, \ldots, x^{N_{T_k}}_{T_k}$. Then, we can write

$$\begin{aligned}
\mathbb{P}(E_{1k} \cap \cdots \cap E_{\kappa k}) &= \int_{(\bar{\mathscr{D}}_G)^\kappa} \mathbb{P}\Big(E_{1k} \cap \cdots \cap E_{\kappa k} \Big| (x^1_{T_k}, \ldots, x^\kappa_{T_k}) \in (\bar{\mathscr{D}}_G)^\kappa\Big) dx^1_{T_k} \cdots dx^\kappa_{T_k} \\
&\leq \int_{(\bar{\mathscr{D}}_G)^\kappa} \mathbb{P}(\Omega_0^{\kappa k}) dx^1_{T_k} \cdots dx^\kappa_{T_k} \\
&= \int_{(\bar{\mathscr{D}}_G)^\kappa} \sum_{\ell=0}^{N_{T_k} - \kappa} \Big((1-p)^\ell \frac{\left(\int_{B(x^1_{T_k}, \ldots, x^\kappa_{T_k})} \lambda_k^\star\right)^\ell}{\ell!} \mathrm{e}^{-\int_{B(x^1_{T_k}, \ldots, x^\kappa_{T_k})} \lambda_k^\star}\Big) \\
&\quad \times \Big(\frac{\left(\int_{\bar{\mathscr{D}}_G \setminus B(x^1_{T_k}, \ldots, x^\kappa_{T_k})} \lambda_k^\star\right)^{N_{T_k} - \kappa - \ell}}{(N_{T_k} - \kappa - \ell)!} \mathrm{e}^{-\int_{\bar{\mathscr{D}}_G \setminus B(x^1_{T_k}, \ldots, x^\kappa_{T_k})} \lambda_k^\star}\Big) dx^1_{T_k} \cdots dx^\kappa_{T_k},
\end{aligned}$$

and the claim follows easily by similar arguments as in the previous proof of Proposition 1. □

The elicitation of the probability of several isolated members is somewhat more difficult than the case of one isolated member. Then, attention shows that the elicitation of the event $\Omega_0^{\kappa k}$ used in the proof of Proposition 2 gives only an upper bound. Unfortunately, this seems to be difficult to prove in a general setting. To establish an asymptotic expression for this probability we need a deeper development and more notations. Let $\mathbb{S}_{r_{\mathsf{af}}}(x^1, \ldots, x^\kappa)$ denote the sub-network over $(x^1, \ldots, x^\kappa)$ in which two vertices are connected (by affinity) if and only if their distance is at most r_{af}. For any integer $\mathbf{n}$ satisfying $1 \leq \mathbf{n} \leq \kappa$, we denote by $\mathcal{C}_{\kappa \mathbf{n}}$ the set of κ-tuples $(x^1, \ldots, x^\kappa) \in (\bar{\mathscr{D}}_G)^\kappa$ satisfying that $\mathbb{S}_{2r_{\mathsf{af}}}(x^1, \ldots, x^\kappa)$ has exactly $\mathbf{n}$ connected components. Notice that the set $\mathcal{C}_{\kappa\kappa}$ consists of those tuples of κ vertices which satisfy, for $i = 1, \ldots, \kappa$, $B(x^i)$ contains none of the other vertices of the tuple. In the following proposition, we state an interesting formula for the computation of the probability of several isolated vertices.

Proposition 3 *Under A3, for any $\kappa \geq 2$ and $(x^1_{T_k}, \ldots, x^\kappa_{T_k}) \in \mathcal{C}_{\kappa\kappa}$, we have for $\alpha, k \to \infty$ and $k \in$*

$\{1,\ldots,k_0\}$

$$\mathbb{P}(E_{1k}\cap\cdots\cap E_{\kappa k}) = \frac{\big(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!\,\mathrm{e}^{\left(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\right)}} \times\int_{\mathcal{C}_{\kappa\kappa}}\Big(1-p\mathrm{vol}(x_{T_k}^1,\ldots,x_{T_k}^\kappa)\Big)^{N_{T_k}-\kappa}dx_{T_k}^1\cdots dx_{T_k}^\kappa,$$

where $\tilde{L}_k\in\mathscr{C}_k$, $\alpha,\beta,A_{\mathsf{f}}>0$, $0<p\leq 1$, $k^{\mathsf{inv}}(\cdot,\cdot)$ *and* $k^{\mathsf{af}}(\cdot)$ *are two probability densities on* $\mathbb{R}^2$ *and* $\bar{\mathscr{D}}_G$.

Proof For any $\kappa\geq 2$ and $(x_{T_k}^1,\ldots,x_{T_k}^\kappa)\in\mathcal{C}_{\kappa\kappa}$, let $\Omega_1^{\kappa k}$ be the event that, for all $1\leq i\leq\kappa$, $B(x_{T_k}^i)$ does not contain any active vertices in states $x_{T_k}^{\kappa+1},\ldots,x_{T_k}^{N_{T_k}}$. In addition, let $\Omega_2^{\kappa k}$ be the event that, for all $1\leq i\leq\kappa$, $B(x_{T_k}^i)$ contains $\mathbf{n}_i$ inactive vertices and no active vertices in states $x_{T_k}^{\kappa+1},\ldots,x_{T_k}^{N_{T_k}}$. Thanks to the previous introducing two events, it follows that

$$\begin{aligned}
\mathbb{P}(E_{1k}\cap\cdots\cap E_{\kappa k}) &= \int_{\mathcal{C}_{\kappa\kappa}}\mathbb{P}\Big(E_{1k}\cap\cdots\cap E_{\kappa k}\Big|(x_{T_k}^1,\ldots,x_{T_k}^\kappa)\in\mathcal{C}_{\kappa\kappa}\Big)dx_{T_k}^1\cdots dx_{T_k}^\kappa\\
&= \int_{\mathcal{C}_{\kappa\kappa}}\mathbb{P}(\Omega_1^{\kappa k})dx_{T_k}^1\cdots dx_{T_k}^\kappa = \int_{\mathcal{C}_{\kappa\kappa}}\sum_{\mathbf{n}_1+\cdots+\mathbf{n}_\kappa}^{N_{T_k}-\kappa}\mathbb{P}(\Omega_2^{\kappa k})dx_{T_k}^1\cdots dx_{T_k}^\kappa\\
&= \int_{\mathcal{C}_{\kappa\kappa}}\sum_{\mathbf{n}_1+\cdots+\mathbf{n}_\kappa=0}^{N_{T_k}-\kappa}\Big\{\prod_{i=1}^{\kappa}\Big((1-p)^{\mathbf{n}_i}\frac{\big(\int_{B(x_{T_k}^i)}\lambda_k^\star\big)^{\mathbf{n}_i}}{\mathbf{n}_i!}\mathrm{e}^{-\int_{B(x_{T_k}^i)}\lambda_k^\star}\Big)\Big\}\\
&\times\Big(\frac{\big(\int_{\bar{\mathscr{D}}_G\setminus B(x_{T_k}^1,\ldots,x_{T_k}^\kappa)}\lambda_k^\star\big)^{N_{T_k}-\kappa-\sum_{i=1}^\kappa\mathbf{n}_i}}{(N_{T_k}-\kappa-\sum_{i=1}^\kappa\mathbf{n}_i)!}\mathrm{e}^{-\int_{\bar{\mathscr{D}}_G\setminus B(x_{T_k}^1,\ldots,x_{T_k}^\kappa)}\lambda_k^\star}\Big)dx_{T_k}^1\cdots dx_{T_k}^\kappa\\
&= \int_{\mathcal{C}_{\kappa\kappa}}\frac{\mathrm{e}^{-\left(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\right)\left(\sum_{i=1}^\kappa\mathrm{vol}(x_{T_k}^i)+1-\mathrm{vol}(x_{T_k}^1,\ldots,x_{T_k}^\kappa)\right)}}{(N_{T_k}-\kappa)!}\\
&\times\sum_{\mathbf{n}_1+\cdots+\mathbf{n}_\kappa=0}^{N_{T_k}-\kappa}\binom{N_{T_k}-\kappa}{\mathbf{n}_1,\ldots,\mathbf{n}_\kappa,N_{T_k}-\kappa-\sum_{i=1}^\kappa\mathbf{n}_i}\\
&\times\Big\{\prod_{i=1}^{\kappa}\Big((1-p)\big(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\big)\big(\mathrm{vol}(x_{T_k}^i)\big)\Big)^{\mathbf{n}_i}\Big\}\\
&\times\Big(\big(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\big)\\
&\times\big(1-\mathrm{vol}(x_{T_k}^1,\ldots,x_{T_k}^\kappa)\big)\Big)^{N_{T_k}-\kappa-\sum_{i=1}^\kappa\mathbf{n}_i}dx_{T_k}^1\cdots dx_{T_k}^\kappa\\
&= \frac{\big(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!\,\mathrm{e}^{\left(\alpha k^{\mathsf{inv}}(\tilde{c},z)+\beta+\frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\right)}}\\
&\qquad\times\int_{\mathcal{C}_{\kappa\kappa}}\Big(1-p\mathrm{vol}(x_{T_k}^1,\ldots,x_{T_k}^\kappa)\Big)^{N_{T_k}-\kappa}dx_{T_k}^1\cdots dx_{T_k}^\kappa,
\end{aligned}$$

where the last line is obtained from the multinomial theorem and by noting that, for any set $(x^1_{T_k},\dots,x^\kappa_{T_k}) \in \mathcal{C}_{\kappa\kappa}$, we have $\sum_{i=1}^\kappa \mathrm{vol}(x^i_{T_k}) = \mathrm{vol}(x^1_{T_k},\dots,x^\kappa_{T_k})$. This completes the proof. □

Our previous statement (Proposition 3) gives an asymptotic equivalence for the probability of several isolated vertices not in the whole domain $(\bar{\mathscr{D}}_G)^\kappa$ but in a more restricted domain given by $(x^1_{T_k},\dots,x^\kappa_{T_k}) \in \mathcal{C}_{\kappa\kappa}$ where $\kappa \geq 2$. Even if this result appears somewhat restrictive, we will show in the sequel that the probability in the domain $(\bar{\mathscr{D}}_G)^\kappa \setminus \mathcal{C}_{\kappa\kappa}$ (which is not covered by our result) converges asymptotically to nothing.

Let us now highlight the major formulas of the previous results. Let denote

$$\alpha^\star_k(N_{T_k}) := \big(\alpha k^{\mathsf{inv}}(\tilde{c},z) + \beta + \frac{1}{2}A_{\mathsf{f}}\tilde{L}_k\gamma_k^2\log(N_{T_k})k^{\mathsf{af}}(y)\big), \tag{3.8}$$

which may be infinite. From Proposition 1, we have asymptotically

$$\mathbb{P}(E_{ik}) \propto \frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)}, \tag{3.9}$$

and we know from Proposition 3 that

$$\mathbb{P}(E_{1k}\cap\cdots\cap E_{\kappa k}) \propto \frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)}. \tag{3.10}$$

The numerical problem that we aim to solve now for establishing the distribution of isolated vertices is the following; We have to constrain $\alpha^\star$ to fulfill the following identity (3.11) needed for the proof of isolated vertices law based on a version of Brun's sieve theorem and Bonferroni inequalities:

$$\frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)} = \Big(\frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)}\Big)^\kappa, \tag{3.11}$$

which is straightforwardly solved by

$$\alpha^\star_k(N_{T_k}) = -N_{T_k}W\Big(-\frac{\big((N_{T_k}-1)!((N_{T_k}-\kappa)!)^{\frac{-1}{\kappa}}\big)^{\frac{\kappa}{(\kappa-1)N_{T_k}}}}{N_{T_k}}\Big), \tag{3.12}$$

where $W(\cdot)$ is the Lambert function. From now on we impose an additional constraint on the constants $\alpha,\beta,A_{\mathsf{f}},\gamma_k$ in order that $\alpha^\star_k$ fulfills (3.11) together with of course the fact that $\alpha,\beta,A_{\mathsf{f}}$ and $\gamma_k > \sqrt{6}/\delta_k$ with $\delta_k \in [0,\delta'_k]$ and $\delta'_k > 0$. One may easily satisfy the constraint (3.11) if one fixes the intrinsic network parameters $\alpha,\beta,A_{\mathsf{f}}$ and one calibrates the needed value of γ_k from formula (3.12).

3.2 Poisson approximation

We now come to our main results. For our social network, the number of isolated vertices (denoted from now on M^k_0) enjoys a Poisson approximation when the size of the network tends to ∞ $\mathbb{P}$-almost surely. So, for $m \in \mathbb{N}$, we have

$$\mathbb{P}\Big(M^k_0 = m\Big) \xrightarrow{\alpha,k\to\infty} \mathrm{e}^{-\mathbb{E}[M^k_0]}\frac{\big(\mathbb{E}[M^k_0]\big)^m}{m!}. \tag{3.13}$$

In this paper we prove result of this kind for the class of measure-valued processes described in Section 2 when we connect each pair of vertices with an indicator function $\mathsf{aff}(\cdot,\cdot)$ of the distance between them. We show that the approximation (3.13) holds for the interacting particle system $\mathscr{N}_{T_k}$ (when $\alpha, k \to \infty$ which guarantees a.s. a large network size), uniformly over affinity functions that are zero beyond a given distance.

We state this more precisely in the next Theorem 2 below, which requires also the Assumption A2. Our main result is as follows (its proof is delayed until Section 5).

Theorem 2 *Assume that A2 and A3 are satisfied. Let*

$$r^{\star}_{\mathsf{af}}(N_{T_k}) = \Big(\frac{\log\Big(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\Big) + \mathfrak{f}_k(N_{T_k}) + (N_{T_k}-1)\log(\alpha^{\star}_k(N_{T_k})) - \alpha^{\star}_k(N_{T_k})}{\pi p N_{T_k}}\Big)^{1/2}, \tag{3.14}$$

for any $p \in (0,1]$ and where $\alpha^{\star}_k$ is given by (3.8) and satisfying (3.11)-(3.12). Then, we have

$$N_{T_k}\Big(\frac{\big(\alpha^{\star}_k(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha^{\star}_k(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x)}dx \longrightarrow \phi, \tag{3.15}$$

as $\alpha, k \to \infty$ and where

$$\phi = \mathrm{e}^{-\lim\limits_{\alpha,k\to\infty}\mathfrak{f}_k(N_{T_k})}.$$

If $\phi \in (0,\infty)$, then as $\alpha, k \to \infty$, we have for $m \in \mathbb{N}$ that

$$\mathbb{P}\Big(M^k_0 = m\Big) \longrightarrow \mathrm{e}^{-\phi}\phi^m/m!. \tag{3.16}$$

If $\phi = 0$, then $\mathbb{P}\Big(M^k_0 = 0\Big) \to 1$, and if $\phi = \infty$, then $\mathbb{P}\Big(M^k_0 = m\Big) \to 0$ for all $m \in \mathbb{N}$.

Indeed, the affinity threshold (3.14) established here for the measure-valued process that generates the data $(\mathbf{x}_k, \mathbf{a}_k)_{1\le k\le k_0}$ looks somewhat complicated than the threshold obtained with uniform random vertices in static geometric graphs. The particularity of the threshold (3.14) is that, in addition to the random size of the data at time $T_k \in [0, T_{k_0}]$, it naturally depends on the intrinsic parameters of the measure-valued process $(\alpha, \beta, A_{\mathsf{f}}, k^{\mathsf{inv}}, k^{\mathsf{af}})$ and on some other parameters $(\gamma_k, \tilde{L}_k)$ needed for discretizing the space into hypercubes of side $r(N_{T_k})/\sqrt{2}$.

We now discuss related work and open problems. Notice that the Poisson distribution (3.13) was already proved by Penrose (2016) [but not with the same threshold (3.14)] in the special case of vertices uniformly distributed respectively in the unit square. Here we are considering a much more general class of measure-valued processes. The result of this paper goes a step beyond the literature in that it considers an interacting particle system with connection (affinity) function that is zero beyond the optimal threshold (3.14) in the same model. As far as the authors know, this is the first work in which results about the optimal threshold and the distribution of isolated vertices are shown for networks with nodes generated by measure-valued processes (rather than the uniform random nodes considered usually in literature). Other ideas for the choose of connection functions are proposed in the literature

that deals with the subject of connectivity in random geometric graphs initiated by Dette & Henze (1989). For instance, using a step connection function ($\mathbb{1}_{[0,\mathsf{r}]}(\|x-y\|)$), Dette & Henze (1989) showed that for any constant $C>0$, the disk graph on n uniform random nodes with connectivity threshold $\mathsf{r}(n)=\sqrt{(\ln(n)+C)/(\pi n)}$ has no isolated nodes with probability $\exp(-e^{-C})$ when n tends to infinity. The theory has been generalized recently in the unit square $[0,1]^d$ (with $d\geq 2$) by Penrose (2016) for a class of connection functions that decay exponentially in some fixed positive power of distance. In some applications, it is desirable to use the Rayleigh fading connection function given by $\exp(-\bar{\beta}(\|x-y\|/\bar{\rho})^{\bar{\gamma}})$ for some fixed positive $\bar{\beta},\bar{\rho},\bar{\gamma}>0$ (typically $\bar{\gamma}=2$). It would be interesting to try to extend our results to these connection functions but this would be a nontrivial task because the discretization method developed here can be quite hard to adaptation.

Another related problem is the connectivity of the network. It is known that the main obstacle to connectivity is the existence of isolated nodes. More clearly, for the geometric (Gilbert) graph $G(\mathcal{X}_n,\mathsf{r}(n))$ with vertex set $\mathcal{X}_n$ given by a set of n independently uniformly distributed nodes in $[0,1]^d$ with $d\geq 2$, and with an edge included between each pair of nodes at distance at most $\mathsf{r}(n)$, Penrose (1997) showed that the probability that the graph is disconnected but free of isolated vertices tends to zero as n tends to infinity, for any choice of $(\mathsf{r}(n))_{n\in\mathbb{N}}$. The same result happens with the Erdös-Rényi graph but the proof for the geometric graph is much harder as pointed out by Bollobás (2001). More formally, for random graphs G, the number of isolated vertices (denoted $M_0(G)$) has asymptotically a Poisson distribution, hence with $\mathcal{K}$ denoting the class of connected graphs we have $\mathbb{P}(G\in\mathcal{K})\sim\mathbb{P}(M_0(G)=0)\sim\exp(-\mathbb{E}[M_0])$ as $n\to\infty$. We would expect something similar to hold for our social network $(\mathscr{N}_t)_{t\geq 0}$. Under additional assumption on $\mathfrak{f}_k$ the connectivity result of graphs G presented here might naturally be conjectured for $(\mathscr{N}_t)_{t\geq 0}$ as follows

$$\mathbb{P}\Big(G_k\big(\mathscr{N}_{T_k},r^{\star}_{\mathsf{af}}(N_{T_k})\big)\in\mathcal{K}\Big)\to e^{-\phi}\qquad\qquad\text{as }\alpha,k\to\infty, \tag{3.17}$$

with $e^{-\phi}$ interpreted as 0 for $\phi=\infty$. More generally, it would be of interest in its own right to extend this to the case of the process studied here $(\mathscr{N}_t)_{t\geq 0}$.

We conclude our main results by checking a statement shown essentially that the number of active isolated vertices (denoted from now on M_a) enjoys also a Poisson approximation but with a slightly different mean. The following theorem states this.

Theorem 3 *Admit assumptions of Theorem 2 and consider the affinity threshold (3.14). If $\phi\in(0,\infty)$, then as $\alpha,k\to\infty$, we have for $m\in\mathbb{N}$ that*

$$\mathbb{P}\Big(M_a^k=m\Big)\longrightarrow \mathrm{e}^{-p\phi}(p\phi)^m/m!, \tag{3.18}$$

where $p\in(0,1]$ and M_a^k denoting the number of active isolated vertices at time T_k. If $\phi=0$, then $\mathbb{P}\Big(M_a^k=0\Big)\to 1$, and if $\phi=\infty$, then $\mathbb{P}\Big(M_a^k=m\Big)\to 0$ for all $m\in\mathbb{N}$.

In Section 5, we will prove Theorems 2 and 3. We finish this section with a short discussion about the limit of $\mathfrak{f}_k(\cdot)$. As we have seen this limit plays a crucial role in determining the Poisson distribution of isolated nodes in the network. From the formula (3.14) of the optimal threshold, we deduce the

following almost surely limits:

$$\lim_{\alpha,k\to\infty} \log\Big(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\Big) = -\infty,\ \lim_{\alpha,k\to\infty} (N_{T_k}-1)\log(\alpha_k^\star(N_{T_k})) = +\infty,$$
$$\text{and } \lim_{\alpha,k\to\infty} -\alpha_k^\star(N_{T_k}) = -\infty.$$

Furthermore, by series expansion we find

$$\lim_{\alpha,k\to\infty} \log\Big(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\Big) + (N_{T_k}-1)\log(\alpha_k^\star(N_{T_k})) - \alpha_k^\star(N_{T_k}) = -\infty,$$

and

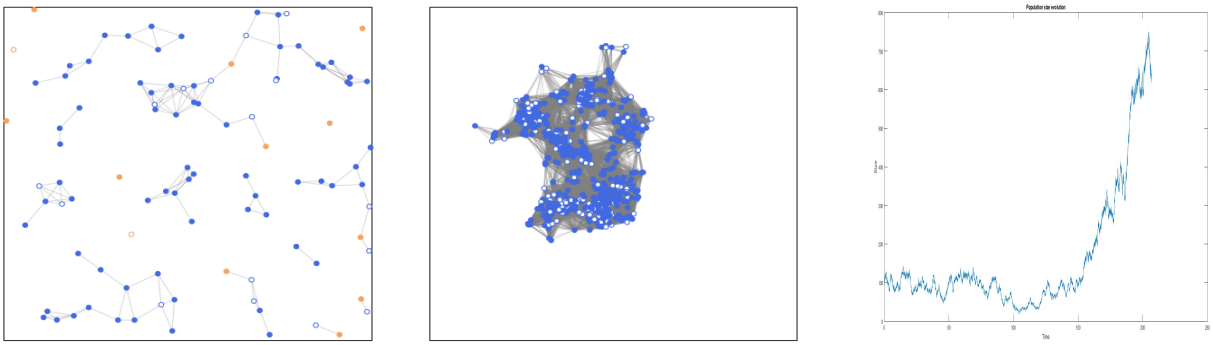
$$\lim_{\alpha,k\to\infty} \frac{\log\Big(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\Big) + (N_{T_k}-1)\log(\alpha_k^\star(N_{T_k})) - \alpha_k^\star(N_{T_k})}{\pi p N_{T_k}} = -\infty.$$

Then, in order that the function ψ_k satisfies the three conditions of Assumption A2, there exists only function $\mathfrak{f}_k(\cdot)$ such that its limit is $+\infty$. The two others cases ($\lim_{\alpha,k\to\infty}\mathfrak{f}_k(N_{T_k}) = \{-\infty, \in (0,\infty)\}$) discussed in Theorems 2 and 3 are in practice impossible to find since there exists no function with these limits together with satisfying at the same time Assumption A2. Hence, ϕ takes only the convenient value 0 and we conclude that there are not any isolated vertices in the graph G_k of the network $\mathscr{N}_{T_k}$ with probability one as $\alpha, k \to \infty$.

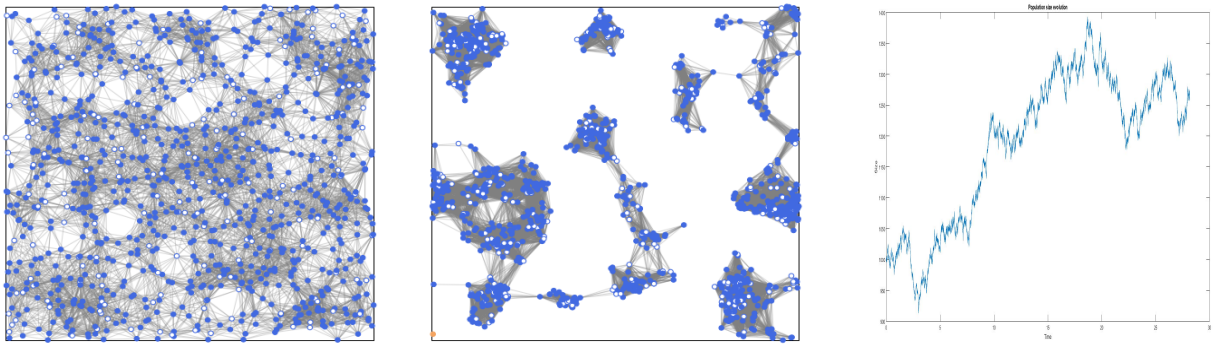
4 Numerical simulations

The simulations in this section were performed using the Monte Carlo algorithm described in Section 2.2. The state space was considered for simplicity $\bar{\mathscr{D}}_G = [0,1]^2$ (supposed being a torus to avoid irregularities around the borders). The initial graph nodes were drawn uniformly on $\bar{\mathscr{D}}_G$. We have chosen for the invitation kernel a bivariate normal distribution $N_2(0,\sigma^2 I_2)$, where $\sigma^2 > 0$ is the variance that determines the dispersion of communities. Concerning the affinity kernel we took a bivariate uniform distribution on $[0,1]^2$, which is somewhat a noninformative choice.

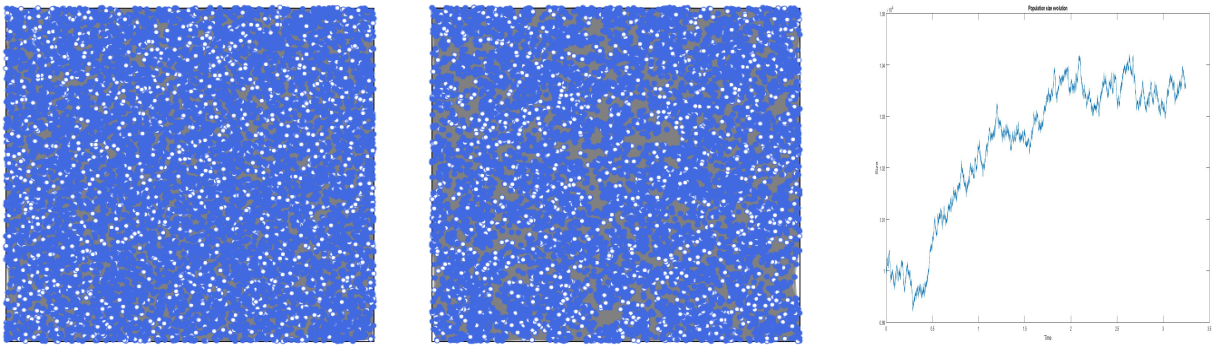
To illustrate the behavior of the dynamic in large graph asymptotic, we performed simulations at different levels of graph size. To put this into practice we vary the number of nodes at the initial time. We considered three cases: (i) $N_0 = 100$, (ii) $N_0 = 1000$ and (iii) $N_0 = 10000$. In each case the Monte Carlo scheme was run for 10^5 updates with parameters $\alpha = 1$, $\beta = 1$, $p = 0.8$, $A_{\mathsf{f}} = 1$, $r_{\mathsf{af}} = 0.1$ and $\sigma = 0.01$. In Figure 2 the evolutions of the graph size and of the graph state are clearly illustrated. In our simulations, the nodes were initially uniformly distributed, but after that communities were formed (see final network states). A certain time, corresponding to the communities formation period, is needed before that the size grows rapidly. This period is shorter when the initial size is bigger. After the formation of communities, the network size grows quickly due to the increase of attractiveness. But as the graph size being bounded, after a long period the network will reach a quasi-stationary state, corresponding to the solution of a deterministic partial differential equation, and the network fluctuates temporarily around this state. The fluctuation can be observed for the two cases with $N_0 = 1000$ and $N_0 = 10000$. The network's extinction is sometimes observed in simulation for the case with $N_0 = 100$.



(a) *Initial graph state with* $N_{T_0} = 100$ (b) *Graph state after* 10^5 *updates* (c) *Node number evolution*



(d) *Initial graph state with* $N_{T_0} = 1000$ (e) *Graph state after* 10^5 *updates* (f) *Node number evolution*



(g) *Initial graph state with* $N_{T_0} = 10000$ (h) *Graph state after* 10^5 *updates* (i) *Node number evolution*

Figure 2: Time evolution of the network state and size during 10^5 Monte Carlo updates starting form small graph size $N_0 = 100$ (above), medium graph size $N_0 = 1000$ (middle) and large graph size $N_0 = 10000$ (bottom). Active nodes are colored in blue, inactive nodes are depicted by empty circles and isolated nodes are colored in orange.

In particular, contrary to the small graph size, the medium and large graph size can not depict the network extinction in the short term where a longer time is needed before the ultimate extinction.

In general, a network is described by the graph and the degree distribution rather than the evolution of its number of nodes. To discuss the effect of affinity function on the behavior of the degree distribution and the graph itself we used the following well-known functions that can be found in the literature to build network models (see, e.g., Aldous (1997); Norros & Reittu (2006); Bollobás et al. (2007); van der Hofstad (2017)):

- The triangular affinity (linear) function given by

$$\mathrm{aff}_2(x,y) = A_{\mathsf{f}}\Big(1 - \frac{1}{r_{\mathsf{af}}}||x-y||\Big)^+, \tag{4.1}$$

  for $x \neq y$. Note that we denote by $(\cdot)^+$ the positive part. Note also that the positive constant $A_{\mathsf{f}} > 0$ is now the maximal affinity that two nodes may reach.

- The logistic affinity function (Kisd's asymmetric function (Kisdi, 1999)) given by

$$\mathsf{aff}_3(x,y) = A_{\mathsf{f}}\Big(\frac{1+c_1}{c_1}\Big)\Big(1 - \frac{1}{1+c_1\exp(-c_2||x-y||)}\Big), \tag{4.2}$$

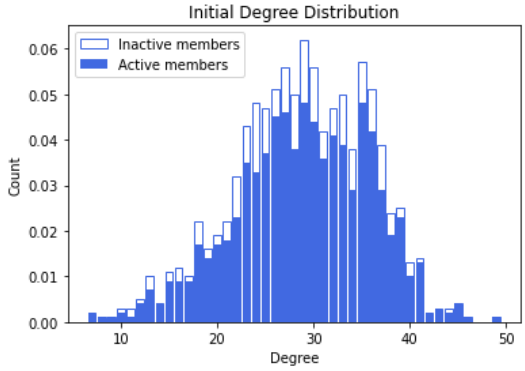
  where c_1 and c_2 are positive constants.

- The Rayleigh fading connection function given by

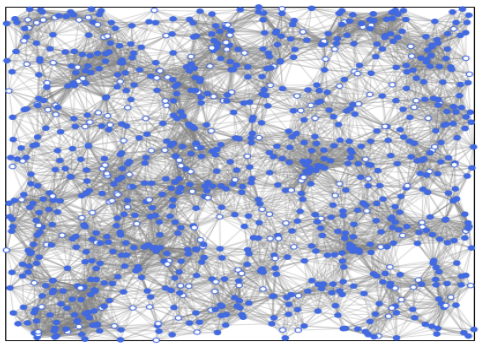
$$\mathsf{aff}_4(x,y) = A_{\mathsf{f}}\exp\Big(-c_3\Big(\frac{||x-y||^{c_5}}{c_4^{c_5}}\Big)\Big), \tag{4.3}$$

  for some fixed positive $c_3, c_4, c_5 > 0$ (typically $c_5 = 2$ to be similar to a Gaussian kernel).

We shall test the impact of these affinity functions (2.2), (4.1), (4.2) and (4.3) on the dynamics. The Monte Carlo scheme was run again for 10^5 updates with the same parameter values as previously: $\alpha = 1$, $\beta = 1$, $p = 0.8$, $A_{\mathsf{f}} = 1$, $r_{\mathsf{af}} = 0.1$ and $\sigma = 0.01$ for each affinity function. Starting from the same initial conditions summarised in Figure 3 with a graph size $N_0 = 1000$, some numerical results obtained from those different affinity functions are displayed in Figure 4. It's clearly observed an increasing (exponential) trend of graph size with the logistic and the Rayleigh fading functions, which is a priori expected from the expression of both functions. Nevertheless, the graph size increases linearly but more rapidly with the rectangular function than with the triangular function as more recruitment by affinity were accepted in the former case. In all cases, it seems that the closer an affinity function is to its maximum (i.e., $\mathsf{aff} \approx A_{\mathsf{f}}$), the bigger is the final network size and the more dispersed are the nodes. In particular, the initial nodes being distributed uniformly on the unit square and the initial degree distribution seems to be close to a Gaussian distribution (see Figure 3). With the triangular function, communities were formed, some of them were related to others, the degree distribution was no more Gaussian but more widespread, particles whose degree is high appeared. Compared to the triangular function, the rectangular function allowed the acceptance of more recruitments by affinity, so the graph size was greater, communities were more connected and the degree distribution has another small bump around 250-300 degrees. With all the four functions, the degree distributions were multimodal and heavy-tailed (there exist many nodes whose degree exceeds the mean degree).



(a) *Degree distribution at time* $t = 0$



(b) *Graph state at time* $t = 0$

Figure 3: Initial degree distribution of active and inactive nodes with the initial network state of $N_0 = 1000$ nodes.

We apply now our main results (Theorems 2 and 3) to the problem of detecting isolated nodes from the random states of the network in a dynamic manner. We explain why our strategy can lead to a perfect analysis especially towards the detection of isolated nodes based on the optimal threshold $r^{\star}_{\mathsf{af}}$ given by (3.14).

The solution to the detection of isolated nodes is typically algorithmic and entirely data based. We try here to offer an answer to this question for the case of our random data $(\mathbf{x}_k, \mathbf{a}_k, N_{T_k})_{1\leq k\leq k_0}$, with k_0 a stopping time, issued from the present network process. To describe the procedure, we begin by defining the function $\mathfrak{f}_k(\cdot)$ for all $k \geq 1$, for instance, by

$$\mathfrak{f}_k(N_{T_k}) := (N_{T_k})^l - \log\left(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\right) - (N_{T_k}-1)\log(\alpha^{\star}_k(N_{T_k})) + \alpha^{\star}_k(N_{T_k}), \tag{4.4}$$

where $l \in (0,1)$. This convenient choice gives immediately $r^{\star}_{\mathsf{af}}(N_{T_k}) = \sqrt{\frac{(N_{T_k})^l}{\pi p N_{T_k}}}$ and, with this particular choice, the threshold $r^{\star}_{\mathsf{af}}$ is not affected by the intrinsic parameters of the network $(\alpha, \beta, A_{\mathsf{f}}, k^{\mathsf{inv}}, k^{\mathsf{af}})$ and the parameters of the discretization (γ_k, L'_k). The algorithm of isolated node detection has three steps, and the main step 2 produces an estimator of one group around one vertex from which we derive the type of nodes (isolated or not). The three steps of the procedure are:

(*i*) Start with an arbitrary starting random vertex $x^i_{T_k}$ where $i \sim U_{\{1,\ldots,N_{T_k}\}}$.

(*ii*) Compute $\mathrm{Card}(x^i_{T_k}) := |B(x^i_{T_k}, r^{\star}_{\mathsf{af}}) \cap \bar{\mathscr{D}}_G|$ and this vertex's $r^{\star}_{\mathsf{af}}$-neighborhood is retrieved, and if it contains, except $x^i_{T_k}$ itself, at most only inactive vertices the vertex is labeled as isolated. Otherwise, the vertex is not isolated. This process continues until the $r^{\star}_{\mathsf{af}}$-neighborhood is completely labeled.

(*iii*) A new unvisited vertex is retrieved and processed, leading to the discovery of a further isolated vertex or not.

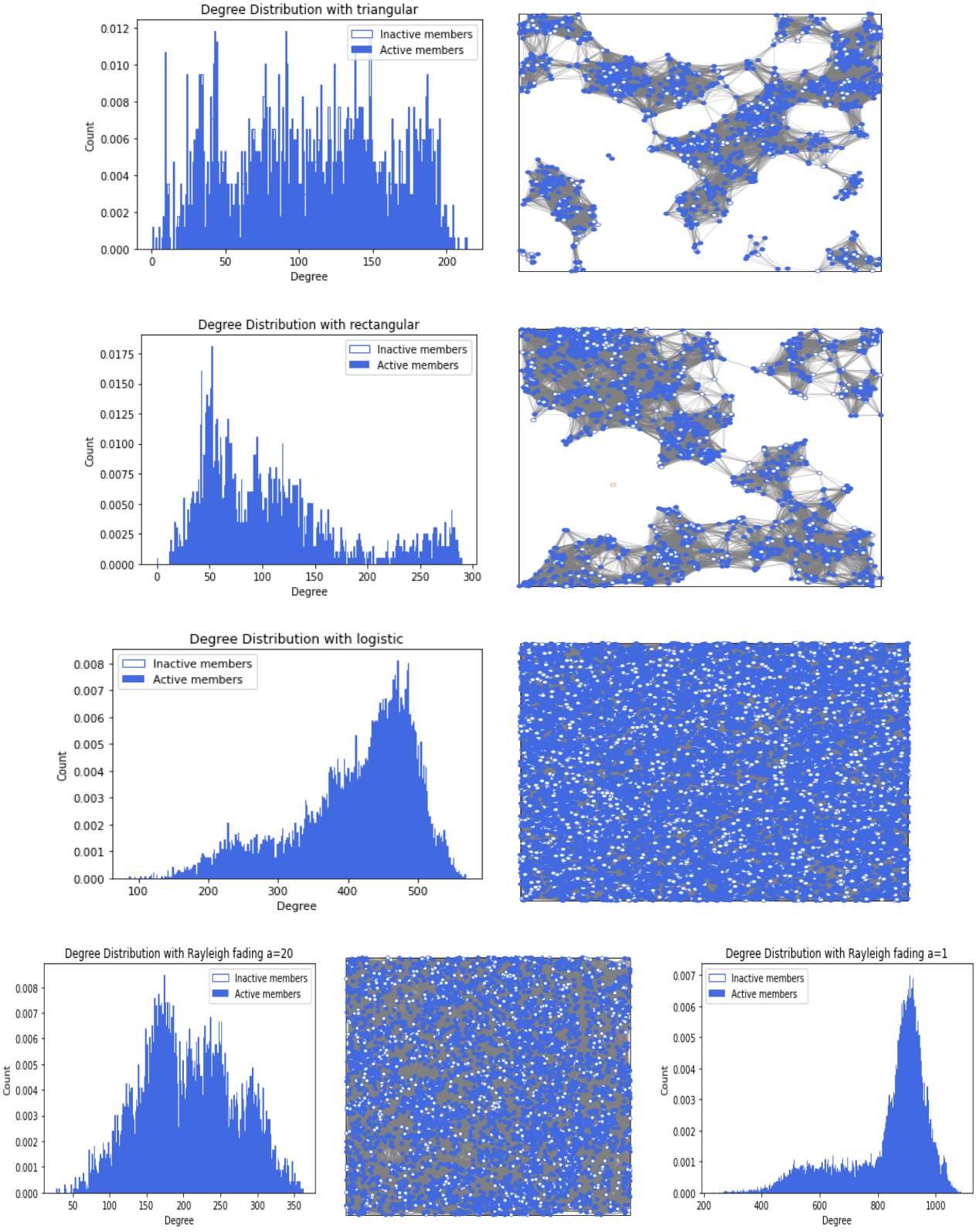
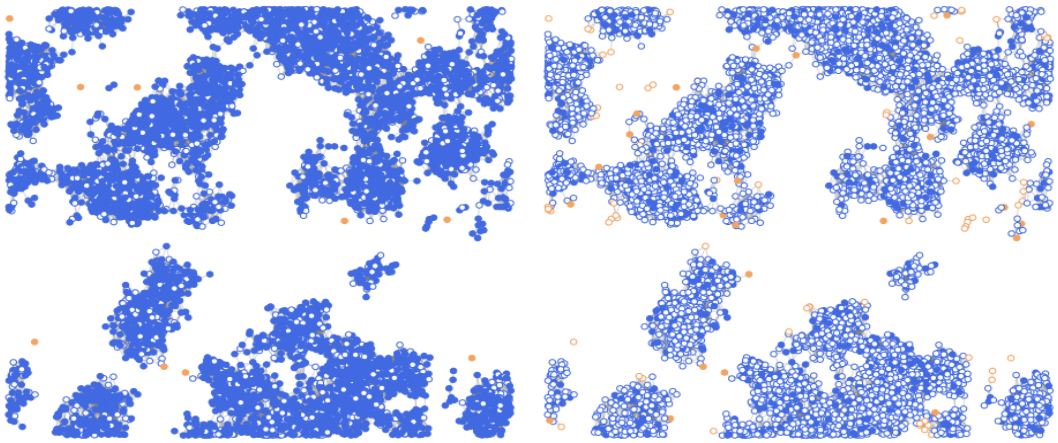


Figure 4: Network degree distribution and state at time $t = T_{10^5}$ starting form graph depicted in Figure 3 with triangular affinity (above), rectangular affinity (middle), logistic affinity (middle) and Rayleigh Fading affinity (bottom). Active nodes are colored in blue, inactive nodes are depicted by empty circles and isolated nodes are colored in orange. Here, we took $c_1 = 1$, $c_2 = 4$, $c_3 = 1, 20$, $c_4 = 1$ and $c_5 = 2$.



(a) *Detection of isolated nodes when* $p = 0.8$ (b) *Detection of isolated nodes when* $p = 0.2$

Figure 5: Figures shown two states of the network (when $p = 0.8$ and $p = 0.2$) at time T_k, here $k = 5 \times 10^4$, simulated with parameters $N_0 = 100$, $\alpha = 2$, $\beta = 1$, $A_{\mathsf{f}} = 1$, $r_{\mathsf{af}} = 0.1$, $\sigma = 0.01$ and their isolated nodes detection analyses.

The construction of an accurate function $\mathfrak{f}_k(\cdot)$ is a crucial step for guaranteeing the statistical optimality of the procedure and for which the results of Theorems 2 and 3 hold. Estimating θ (if it is unknown) before estimating the partition itself is a non-trivial task, and needs to be done with extreme care. Note that the constant $l \in (0, 1)$ is completely determined when one assesses the model parameters r_{af} and p. The required inputs for Step 2 of the procedure are: (i) N_{T_k}, the current size of the network; (ii) p, the probability to be active in the network. Hence, this step is done at no additional accuracy cost. Remark that, unlikely a K-means procedure, the detection procedure considered here does not require the number of isolated nodes, which needs an approach for selecting it in a data-adaptive fashion.

Figure 5 shows an application of the isolated nodes detection analysis of the network at time T_k, here $k = 5 \times 10^4$. The size of this network at this time is $N_{T_{50000}} = 12646$ nodes. For the isolated nodes detection shown in Figure 5(a), the function (4.4) was used with $l = \log((r_{\mathsf{af}})^2 p\pi N_{T_k})/\log(N_{T_k})$ where $p = 0.8$ and $r_{\mathsf{af}} = 0.1$. This gives here $l = 0.61$. The number of isolated vertices founded by the isolated nodes detection procedure was 10. To compare this analysis with a less probability of activity, we plot in Figure 5(b) the same procedure applied to the same time but now with $p = 0.2$. The size of the network at this time was $N_{T_{50000}} = 11952$ nodes. We obtained $l \approx 0.46$, and in this case we observed 96 isolated nodes. Hence, a small probability of activity p induces as one can predict more isolated nodes in the network.

Finally, we increased the invitation rate to $\alpha = 10000$ in order to strengthen the occupation of the unit square by the network nodes. We concluded by observing (Figure not given here) that the isolated nodes have disappeared and the network forms one connected group which confirms, even numerically, our previous conjecture on the connectivity of the network $[\mathbb{P}\big(G_k\big(\mathscr{N}_{T_k}, r^{\star}_{\mathsf{af}}(N_{T_k})\big) \in \mathcal{K}\big) \to e^{-\phi} = 1$ as $\alpha, k \to \infty]$.

5 Proofs of main results

In this section we prove Theorems 2 and 3 where we propose to split the proof into several lemmas. Our strategy of proof is as follows. We partition the state space $\bar{\mathscr{D}}_G$ into four regions $\mathcal{D}_1, \ldots, \mathcal{D}_4$ as explained in Figure 6. From this partition, it is easy to see that for any x in $\bar{\mathscr{D}}_G$ we have $\pi(r_{\mathsf{af}})^2/4 \leq \mathrm{vol}(x) \leq \pi(r_{\mathsf{af}})^2$ (in particular $\mathrm{vol}(x) = \pi(r_{\mathsf{af}})^2/4$ when x is at the corners and $\mathrm{vol}(x) = \pi(r_{\mathsf{af}})^2$ for all $x \in \mathcal{D}_1$).

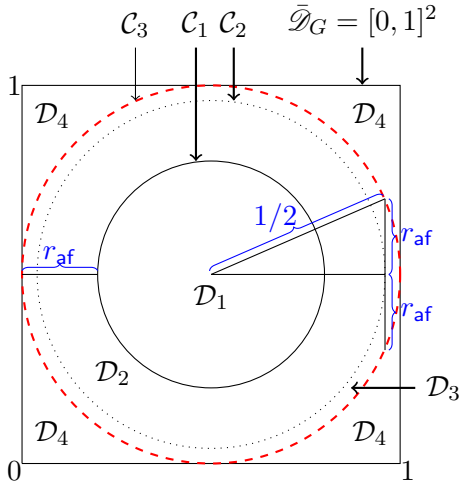


Figure 6: The different levels of $\bar{\mathscr{D}}_G$ partitioning are shown here with three circles $\mathcal{C}_1$ (–), $\mathcal{C}_2$ ($\cdots$) and $\mathcal{C}_3$ (- -). The space is subdivided into 4 regions; $\mathcal{D}_1$ (given by the disk of radius $1/2 - r_{\mathsf{af}}$ centered at $\mathbb{O} = (1/2, 1/2)$), $\mathcal{D}_2$ (given by the annulus of radii $1/2 - r_{\mathsf{af}}$ and $\sqrt{1/4 - (r_{\mathsf{af}})^2}$ centered at $\mathbb{O}$), $\mathcal{D}_3$ (given by the annulus of radii $\sqrt{1/4 - (r_{\mathsf{af}})^2}$ and $1/2$ centered at $\mathbb{O}$) and $\mathcal{D}_4$ (given by the rest of $\bar{\mathscr{D}}_G$).

In the following lemma, we shall study a lower bound on $\mathrm{vol}(x)$ for all $x \in \mathcal{D}_2$.

Lemma 3 *Under the previous partition of $\bar{\mathscr{D}}_G$ given in Figure 6, then, for all $x \in \mathcal{D}_2$, we have*

$$\mathrm{vol}(x) \geq \frac{\pi(r_{\mathsf{af}})^2}{2} + r_{\mathsf{af}}\Big(\frac{1}{2} - \eta\|x\|\Big), \qquad \textit{with} \qquad \eta \geq \frac{1}{\frac{\sqrt{1/2 - 2(r_{\mathsf{af}})^2}}{1/2 - 2(r_{\mathsf{af}})^2} - 1}, \tag{5.1}$$

where $0 < r_{\mathsf{af}} \leq 1/2$.

Proof Without loss of generality, let consider two vertices $x, y \in \bar{\mathscr{D}}_G$ such that y be a vertex in $\partial\bar{\mathscr{D}}_G$ (the boundary of $\bar{\mathscr{D}}_G$) with coordinates $(0.5, 0)$ or $(1, 0.5)$ or $(0.5, 1)$ or $(0, 0.5)$. Straightforwardly, at these fourth coordinates the distance between the circle $\mathcal{C}_2$ and the boundary of $\bar{\mathscr{D}}_G$ achieves its minimum. It is sufficient to prove the lemma for one of these coordinates. The vertex x be in $\partial\mathcal{C}_2$ together with $\|y - x\| = 1/2 - \|x - \mathbb{O}\|$ and ab be the diameter of $B(x, r_{\mathsf{af}})$ perpendicular to xy (see Figure 7).

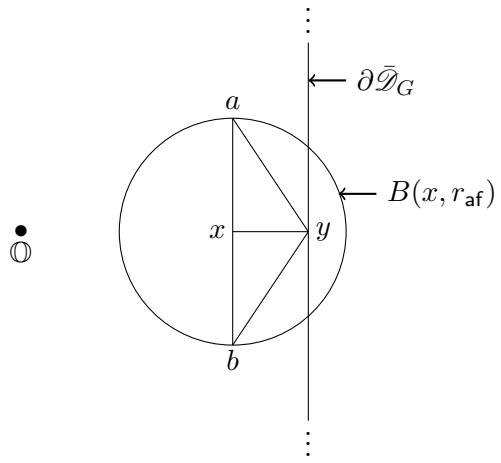


Figure 7: A partial configuration of $\bar{\mathscr{D}}_G$ boundary and the ball $B(x, r_{\mathsf{af}})$ that contains the vertex $y = (1, 0.5)$ and centered at $x = (0.5 + \sqrt{1/4 - (r_{\mathsf{af}})^2}, 0.5)$. The volume $\mathrm{vol}(x)$ is higher than the volume of the half of $B(x, r_{\mathsf{af}})$ plus the area of the triangle aby.

Thus, for all $x \in \mathcal{D}_2$, we can write

$$\mathrm{vol}(x) \geq \frac{\pi(r_{\mathsf{af}})^2}{2} + r_{\mathsf{af}}\|y - x\| = \frac{\pi(r_{\mathsf{af}})^2}{2} + r_{\mathsf{af}}\Big(1/2 - \|x - \mathbb{O}\|\Big). \tag{5.2}$$

Now, for all $x \in \partial\mathcal{C}_2$ and from Figure 6, we have

$$\inf_{x \in \partial\mathcal{C}_2} \|x\| = \frac{\sqrt{2}}{2} - \sqrt{1/4 - (r_{\mathsf{af}})^2} = \frac{\sqrt{2}}{2} - \|x - \mathbb{O}\| \leq \|x\|.$$

One may easily find an upper bound for $\|x - \mathbb{O}\|$ in function of $\|x\|$ by assuming that exists $\eta \in (0, \infty)$ such that $\|x - \mathbb{O}\| \leq \eta \inf_{x \in \partial\mathcal{C}_2} \|x\|$ and by easy calculation we find

$$\begin{aligned} \|x - \mathbb{O}\| \leq \eta\Big(\frac{\sqrt{2}}{2} - \|x - \mathbb{O}\|\Big) \qquad &\Leftrightarrow \qquad \sqrt{1/4 - (r_{\mathsf{af}})^2} \leq \eta\Big(\frac{\sqrt{2}}{2} - \sqrt{1/4 - (r_{\mathsf{af}})^2}\Big) \\ &\Leftrightarrow \eta \geq \frac{1}{\frac{\sqrt{1/2 - 2(r_{\mathsf{af}})^2}}{1/2 - 2(r_{\mathsf{af}})^2} - 1}. \end{aligned}$$

Hence, we complete the proof by using $\|x - \mathbb{O}\| \leq \eta \inf_{x \in \partial\mathcal{C}_2} \|x\| \leq \eta\|x\|$ in (5.2) which leads to (5.1). □

We may now formulate a lower bound on the volume of more than one vertex inside the network, for which the proof follows the same geometric spirit of Lemma 3 proof.

Lemma 4 *Admit the partition of the torus $\bar{\mathscr{D}}_G$ given in Figure 6. For any sequence $(x^1, \ldots, x^\kappa) \in (\bar{\mathscr{D}}_G)^\kappa$ where $\kappa \geq 2$ and such that x^1 has the largest norm with $\|x^i - x^j\| \leq 2r_{\mathsf{af}}$ if and only if $|i - j| \leq 1$, we*

have

$$\mathrm{vol}(x^1,\dots,x^\kappa) \geq \mathrm{vol}(x^1) + \frac{\pi r_{\mathsf{af}}}{16}\sum_{i=1}^{\kappa-1}\Delta x^i, \tag{5.3}$$

where $\Delta x^i = \|x^{i+1} - x^i\|$.

Proof We show the result (5.3) for $\kappa = 2$ at first. Then, we prove by induction that the result holds for any $\kappa \geq 2$. For simplicity, let fix $\kappa = 2$ and consider the function $\varphi(\Delta x^1) = \mathrm{vol}\big(B(x^2, r_{\mathsf{af}}) \setminus B(x^1, r_{\mathsf{af}})\big)$. Our aim is to prove that $\varphi(\Delta x^1) \geq \frac{\pi r_{\mathsf{af}}}{2}\Delta x^1$ for any $x^1, x^2 \in \bar{\mathscr{D}}_G$ with $\|x^1\| \geq \|x^2\|$. To do this, let y_1y_2 be the common chord of $\partial B(x^1, r_{\mathsf{af}})$ and $\partial B(x^2, r_{\mathsf{af}})$. Let also z_1z_2 be another chord of $\partial B(x^2, r_{\mathsf{af}})$ that is parallel to y_1y_2 and has the same length as y_1y_2 as explained in Figure 8.

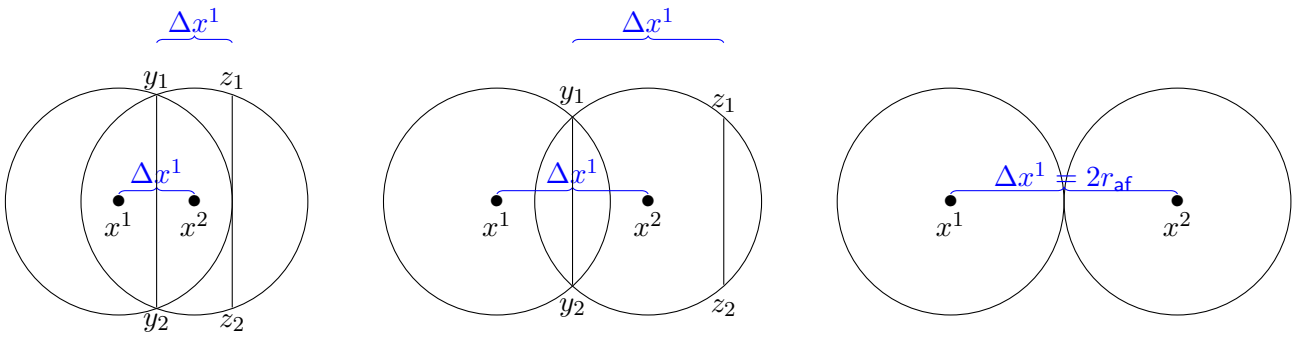


Figure 8: Three examples of configuration for the chords y_1y_2 and z_1z_2 (of two intersecting balls $B(x^1, r_{\mathsf{af}})$ and $B(x^2, r_{\mathsf{af}})$), where we see that $\|y_1 - y_2\|$ is decreasing over $[0, 2r_{\mathsf{af}}]$.

From Figure 8, its clear that $\varphi(\Delta x^1)$ is equal to the volume of the portion of $B(x^2, r_{\mathsf{af}})$ between the chords y_1y_2 and z_1z_2 and immediately we deduce that the second derivative $\varphi''(\Delta x^1) \leq 0$ (since $\varphi'(\Delta x^1) = \|y_1 - y_2\|$ which is decreasing). Hence, the function φ is concave with $\varphi(0) = 0$ and $\varphi(2r_{\mathsf{af}}) = \pi(r_{\mathsf{af}})^2$ which enables us, for $\forall \Delta x, \Delta x' \in [0, 2r_{\mathsf{af}}]$ and $\forall \ell^0 \in [0,1]$, to write

$$\varphi\big(\ell^0\Delta x + (1-\ell^0)\Delta x'\big) \geq \ell^0\varphi(\Delta x) + (1-\ell^0)\varphi(\Delta x'),$$

and by taking $\Delta x = 0$ and $\Delta x' = 2r_{\mathsf{af}}$,

$$\varphi\big((1-\ell^0)2r_{\mathsf{af}}\big) \geq (1-\ell^0)\pi(r_{\mathsf{af}})^2, \tag{5.4}$$

where $\varphi(\Delta x^1) \geq \frac{\pi r_{\mathsf{af}}}{2}\Delta x^1$ holds by choosing ℓ^0 in (5.4) such that $\Delta x^1 = (1-\ell^0)2r_{\mathsf{af}}$. We discuss now the lower bound of $\mathrm{vol}(x^1, x^2) - \mathrm{vol}(x^1)$ following the position of x^1 in the partition of $\bar{\mathscr{D}}_G$. Indeed, if $x^1 \in \mathcal{D}_1$, then the two balls $B(x^1, r_{\mathsf{af}})$ and $B(x^2, r_{\mathsf{af}})$ are completely contained in $\bar{\mathscr{D}}_G$ which enables us to obtain

$$\mathrm{vol}(x^1, x^2) - \mathrm{vol}(x^1) = \mathrm{vol}\big(B(x^2, r_{\mathsf{af}}) \setminus B(x^1, r_{\mathsf{af}})\big) = \varphi(\Delta x^1) \geq \frac{\pi r_{\mathsf{af}}}{2}\Delta x^1,$$

and the lemma is proved for $\kappa = 2$ and if $x^1 \in \mathcal{D}_1$. If $x^1 \notin \mathcal{D}_1$, we notice that for the same distance Δx^1, the value of $\mathrm{vol}(x^1, x^2) - \mathrm{vol}(x^1)$ achieves its minimum if $x^1, x^2 \in \partial\bar{\mathscr{D}}_G$ together with x^2 is at the corner. It suffice to show the lemma in this case. For the easy of exposition, let consider again the previous chords such that $y_2 \in \bar{\mathscr{D}}_G$ as shown in Figure 9.

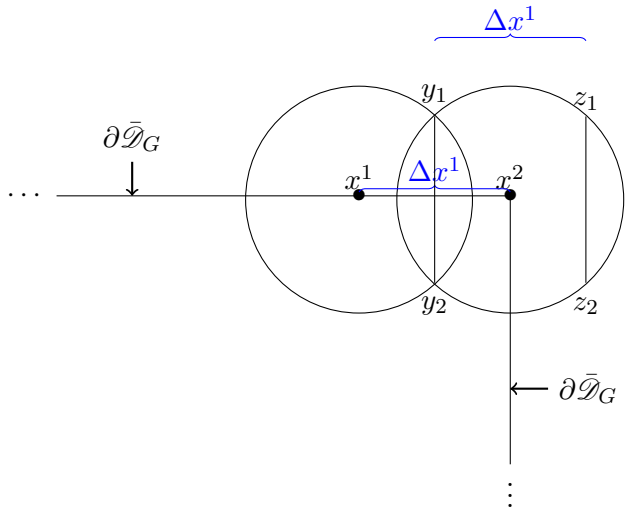


Figure 9: An example of configuration for the chords y_1y_2 and z_1z_2 (of two intersecting balls $B(x^1, r_{\mathsf{af}})$ and $B(x^2, r_{\mathsf{af}})$), where x^2 is at the corner of $\bar{\mathscr{D}}_G$.

From Figure 9, it appeared that $\mathrm{vol}(x^1, x^2) - \mathrm{vol}(x^1) = \frac{\varphi(\Delta x^1)}{4} \geq \frac{\pi r_{\mathsf{af}}}{8}\Delta x^1$ and the lemma follows also in this case. Now, let consider $\kappa = 3$ and write

$$\begin{aligned}\mathrm{vol}(x^1, x^2, x^3) &\geq \mathrm{vol}(x^1) + \mathrm{vol}(x^3) \geq \mathrm{vol}(x^1) + \frac{\pi(r_{\mathsf{af}})^2}{4} \\ &= \mathrm{vol}(x^1) + \frac{\pi r_{\mathsf{af}}}{16} 4 r_{\mathsf{af}} \geq \mathrm{vol}(x^1) + \frac{\pi r_{\mathsf{af}}}{16} \sum_{i=1}^{2} \Delta x^i,\end{aligned}$$

since $\Delta x^i \leq 2r_{\mathsf{af}}$. Finally, we generalize for $\kappa > 3$ by induction as follows

$$\begin{aligned}\mathrm{vol}(x^1, \ldots, x^\kappa) &\geq \mathrm{vol}(x^1, \ldots, x^{\kappa-2}) + \mathrm{vol}(x^\kappa) \\ &\geq \mathrm{vol}(x^1) + \Big(\frac{\pi r_{\mathsf{af}}}{16} \sum_{i=1}^{\kappa-3} \Delta x^i\Big) + \frac{\pi(r_{\mathsf{af}})^2}{4} \\ &\geq \mathrm{vol}(x^1) + \frac{\pi r_{\mathsf{af}}}{16} \sum_{i=1}^{\kappa-1} \Delta x^i,\end{aligned}$$

which completes the proof. □

We prove now a lower bound on the volume of more than one vertex inside the sub-network $\mathbb{S}_{2r_{\mathsf{af}}}(x^1, \ldots, x^\kappa)$ over $(x^1, \ldots, x^\kappa) \in \mathcal{C}_{\kappa 1}$ in which two vertices are connected (by affinity) if and only if their distance is at most $2r_{\mathsf{af}}$.

Lemma 5 *For any sequence $(x^1,\dots,x^\kappa)\in\mathcal{C}_{\kappa 1}$ where $\kappa\geq 2$ and such that x^1 being the one of the largest norm among $x^1,\dots,x^\kappa$, we have*

$$\mathrm{vol}(x^1,\dots,x^\kappa)\geq \mathrm{vol}(x^1)+\frac{\pi r_{\mathsf{af}}}{16}\sup_{2\leq i\leq\kappa}\|x^i-x^1\|. \tag{5.5}$$

Proof For the easy of the proof, we assume that $\|x^\kappa-x^1\|=\sup_{2\leq i\leq\kappa}\|x^i-x^1\|$ and let $\mathcal{P}$ be a minimum-hop path between x^1 and x^κ in $\mathbb{S}_{2r_{\mathsf{af}}}(x^1,\dots,x^\kappa)$ with $\Delta^\star x$ be the total length of $\mathcal{P}$. Thus, every pair of vertices in $\mathcal{P}$ that are not adjacent vertices in $\mathcal{P}$ are distant by more than $2r_{\mathsf{af}}$ and by application of Lemma 4 to the vertices in the path $\mathcal{P}$ we find that

$$\mathrm{vol}\big(\{x^i: x^i\in\mathcal{P}\}\big)\geq \mathrm{vol}(x^1)+\frac{\pi r_{\mathsf{af}}}{16}\Delta^\star x.$$

We conclude by remarking that $\mathrm{vol}(x^1,\dots,x^\kappa)\geq\mathrm{vol}(\{x^i: x^i\in\mathcal{P}\})$ and $\Delta^\star x\geq\|x^\kappa-x^1\|$. This completes the proof. □

In proving (3.15), we shall use the following lemma.

Lemma 6 *Assume that A2 and A3 are satisfied. Let*

$$r^\star_{\mathsf{af}}(N_{T_k})=\Big(\frac{\log\Big(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\Big)+\mathfrak{f}_k(N_{T_k})+(N_{T_k}-1)\log(\alpha^\star_k(N_{T_k}))-\alpha^\star_k(N_{T_k})}{\pi p N_{T_k}}\Big)^{1/2},$$

for any $p\in(0,1]$ and where $\alpha^\star_k$ is given by (3.8). Then, we have

$$N_{T_k}\Big(\frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x)}dx\longrightarrow \mathrm{e}^{-\lim\limits_{\alpha,k\to\infty}\mathfrak{f}_k(N_{T_k})},\qquad \text{as } \alpha,k\to\infty.$$

Proof We shall proceed by approximating the integral in the left-hand side using the four regions $\mathcal{D}_1,\dots,\mathcal{D}_4$ of the unit square $\bar{\mathscr{D}}_G$. If $x\in\mathcal{D}_1$, we know that $\mathrm{vol}(x)=\pi(r^\star_{\mathsf{af}})^2$ and it's straightforward that

$$\begin{aligned}
N_{T_k}\Big(\frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)}\Big)\int_{\mathcal{D}_1}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x)}dx &= N_{T_k}\mathrm{e}^{-\log\left(\frac{N_{T_k}\pi}{4}\right)-\mathfrak{f}_k(N_{T_k})}\int_{\mathcal{D}_1}dx\\
&=\frac{4}{\pi}\mathrm{e}^{-\mathfrak{f}_k(N_{T_k})}\pi\Big(\frac{1}{2}-r^\star_{\mathsf{af}}\Big)^2\\
&=\mathrm{e}^{-\mathfrak{f}_k(N_{T_k})}\Big(1-2r^\star_{\mathsf{af}}\Big)^2\\
&\overset{\alpha,k\to+\infty}{\longrightarrow}\mathrm{e}^{-\lim\limits_{\alpha,k\to\infty}\mathfrak{f}_k(N_{T_k})},
\end{aligned}$$

where we note that for a.s. sufficiently large N_{T_k} we have $r^\star_{\mathsf{af}}\to 0$ (by Assumption A2). Therefore, if $x\in\mathcal{D}_3$, we know that $\mathrm{vol}(x)\geq\frac{1}{2}\pi(r^\star_{\mathsf{af}})^2$ and by using the upper bound

$$\forall k\geq 1,\qquad \Big(\frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha^\star_k(N_{T_k})\right)}\Big)\leq 1,$$

we find

$$N_{T_k}\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{\mathcal{D}_3}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x)}dx \leq N_{T_k}\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\int_{\mathcal{D}_3}dx$$
$$= N_{T_k}\pi(r_{\mathsf{af}}^\star)^2\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\overset{\alpha,k\to+\infty}{\longrightarrow}0.$$

The same think happens if $x\in\mathcal{D}_4$ where $\mathrm{vol}(x)\geq\frac{1}{4}\pi(r_{\mathsf{af}}^\star)^2$ but we need more notation. Let consider the vertices $A=(0,0), B=(\frac{1}{2},0), C=(0,\frac{1}{2})$ and the triangle $\mathbb{T}$ formed by ABC. From Figure 6, it's clear that $\mathrm{vol}(\mathcal{D}_4)\leq 4\mathrm{vol}(\mathbb{T})=\frac{4}{8}$ and we find

$$N_{T_k}\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{\mathcal{D}_4}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x)}dx \leq N_{T_k}\mathrm{e}^{-\frac{1}{4}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\int_{\mathcal{D}_4}dx$$
$$\leq \frac{N_{T_k}}{2}\mathrm{e}^{-\frac{1}{4}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\overset{\alpha,k\to+\infty}{\longrightarrow}0.$$

Let us now turn out to the region $x\in\mathcal{D}_2$. Hence, by application of Lemma 3 and the polar coordinate system we have

$$\begin{aligned}
&N_{T_k}\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{\mathcal{D}_2}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x)}dx\\
&\qquad\leq N_{T_k}\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\int_{\mathcal{D}_2}\mathrm{e}^{-N_{T_k}pr_{\mathsf{af}}^\star\left(\frac{1}{2}-\eta\|x\|\right)}dx\\
&\qquad= 2\pi N_{T_k}\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\int_{\frac{1}{2}-r_{\mathsf{af}}^\star}^{\sqrt{\frac{1}{4}-(r_{\mathsf{af}}^\star)^2}}\rho_x\mathrm{e}^{-N_{T_k}pr_{\mathsf{af}}^\star\left(\frac{1}{2}-\eta\rho_x\right)}d\rho_x\\
&\qquad\leq 2\sqrt{2}\pi N_{T_k}\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\int_{\frac{1}{2}-r_{\mathsf{af}}^\star}^{\frac{1}{2}}\mathrm{e}^{-N_{T_k}pr_{\mathsf{af}}^\star\left(\frac{1}{2}-\eta\rho_x\right)}d\rho_x\\
&\qquad= \frac{2\sqrt{2}\pi}{\eta}N_{T_k}\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\int_{\frac{1}{2}-\frac{1}{2}\eta}^{\frac{1}{2}-\eta(\frac{1}{2}-r_{\mathsf{af}}^\star)}\mathrm{e}^{-N_{T_k}pr_{\mathsf{af}}^\star z}dz\\
&\qquad= \frac{2\sqrt{2}\pi}{\eta pr_{\mathsf{af}}^\star}\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}\Big(1-\mathrm{e}^{-N_{T_k}p\eta(r_{\mathsf{af}}^\star)^2}\Big)\\
&\qquad\leq \frac{2\sqrt{2}\pi}{p}\frac{1}{r_{\mathsf{af}}^\star}\Big(\frac{\sqrt{\frac{1}{2}-2(r_{\mathsf{af}}^\star)^2}}{\frac{1}{2}-2(r_{\mathsf{af}}^\star)^2}-1\Big)\mathrm{e}^{-\frac{1}{2}N_{T_k}p\pi(r_{\mathsf{af}}^\star)^2}=\frac{O(1)}{\sqrt{\psi_k(N_{T_k})}}=o(1),
\end{aligned}$$

by our choice of $\mathfrak{f}_k$, this shows ψ_k tends to infinity, completing the proof. □

We need also the following lemma.

Lemma 7 *Assume that A2 and A3 are satisfied. Let*

$$r_{\mathsf{af}}^\star(N_{T_k})=\Big(\frac{\log\left(\frac{N_{T_k}\pi}{4(N_{T_k}-1)!}\right)+\mathfrak{f}_k(N_{T_k})+(N_{T_k}-1)\log(\alpha_k^\star(N_{T_k}))-\alpha_k^\star(N_{T_k})}{\pi pN_{T_k}}\Big)^{1/2},$$

for any $p \in (0,1]$ *and where* $\alpha_k^\star$ *is given by (3.8) and satisfying (3.11). Then, for any* $\kappa \geq 2$ *and* $(x^1,\dots,x^\kappa) \in \mathcal{C}_{\kappa\mathbf{n}}$ *with* $1 \leq \mathbf{n} < \kappa$*, we have*

$$(N_{T_k})^\kappa \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big) \int_{\mathcal{C}_{\kappa\mathbf{n}}} \mathrm{e}^{-N_{T_k} p \mathrm{vol}(x^1,\dots,x^\kappa)} dx^1 \cdots dx^\kappa \longrightarrow 0,$$

as $\alpha, k \to \infty$*.*

Proof We divide the proof into two steps.

Step 1. Let us first show the lemma for $\mathbf{n} = 1$. For simplicity and without loss of generality, let consider the subset $\mathcal{C}_{\kappa 1}^0$ denoting the set of $(x^1,\dots,x^\kappa) \in \mathcal{C}_{\kappa 1}$ satisfying that x^1 being the one of the largest norm among $x^1,\dots,x^\kappa$ and x^κ being the one with longest distance from x^1 among $x^1,\dots,x^\kappa$ which enables us to write

$$\begin{aligned}
&(N_{T_k})^\kappa \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big) \int_{\mathcal{C}_{\kappa 1}} \mathrm{e}^{-N_{T_k} p \mathrm{vol}(x^1,\dots,x^\kappa)} dx^1 \cdots dx^\kappa \\
&\quad \leq \kappa(\kappa-1)(N_{T_k})^\kappa \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big) \int_{\mathcal{C}_{\kappa 1}^0} \mathrm{e}^{-N_{T_k} p \mathrm{vol}(x^1,\dots,x^\kappa)} dx^1 \cdots dx^\kappa.
\end{aligned}$$

We claim next that (5.5) (in Lemma 5) holds also for $(x^1,\dots,x^\kappa) \in \mathcal{C}_{\kappa 1}^0$ with $\frac{\pi}{16}$ replaced by some constant $\tilde{C}$, that is,

$$\mathrm{vol}(x^1,\dots,x^\kappa) \geq \mathrm{vol}(x^1) + \tilde{C} r_{\mathsf{af}}^\star \|x^\kappa - x^1\|, \tag{5.6}$$

with $x^\kappa \in B(x^1, 2(\kappa-1)r_{\mathsf{af}}^\star)$ and, for $i \in \{2,\dots,\kappa-1\}$, $x^i \in B(x^1, \|x^\kappa - x^1\|)$. Indeed, by the constraint (3.11) imposed on $\alpha^\star$ and the fact that

$$\forall k \geq 1, \qquad \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!} e^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big)^{(\kappa-1)} \leq 1,$$

we obtain

$$
\begin{aligned}
&(N_{T_k})^{\kappa}\Big(\frac{\big(\alpha_k^{\star}(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\mathcal{C}^0_{\kappa 1}}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1,\ldots,x^{\kappa})}dx^1\cdots dx^{\kappa}\\
&\quad\leq (N_{T_k})^{\kappa}\Big(\frac{\big(\alpha_k^{\star}(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\mathcal{C}^0_{\kappa 1}}\mathrm{e}^{-N_{T_k}p\Big(\mathrm{vol}(x^1)+\tilde{C}r^{\star}_{\mathsf{af}}\|x^{\kappa}-x^1\|\Big)}dx^1\cdots dx^{\kappa}\\
&\quad\leq (N_{T_k})^{\kappa}\Big(\frac{\big(\alpha_k^{\star}(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1)}dx^1\\
&\qquad\times\int_{B(x^1,2(\kappa-1)r^{\star}_{\mathsf{af}})}\mathrm{e}^{-N_{T_k}p\tilde{C}r^{\star}_{\mathsf{af}}\|x^{\kappa}-x^1\|}dx^{\kappa}\prod_{i=2}^{\kappa-1}\int_{B(x^1,\|x^{\kappa}-x^1\|)}dx^i\\
&\quad= 2(\pi N_{T_k})^{\kappa-1}\Big(N_{T_k}\Big(\frac{\big(\alpha_k^{\star}(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1)}dx^1\Big)\\
&\qquad\times\Big(\int_0^{2(\kappa-1)r^{\star}_{\mathsf{af}}}(\rho_x)^{2\kappa-3}\mathrm{e}^{-N_{T_k}p\tilde{C}r^{\star}_{\mathsf{af}}\rho_x}d\rho_x\Big)
\end{aligned}
$$

$$
\begin{aligned}
&< 2(\pi N_{T_k})^{\kappa-1}\Big(N_{T_k}\Big(\frac{\big(\alpha_k^{\star}(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1)}dx^1\Big)\\
&\qquad\times\Big(\int_0^{\infty}(\rho_x)^{2\kappa-3}\mathrm{e}^{-N_{T_k}p\tilde{C}r^{\star}_{\mathsf{af}}\rho_x}d\rho_x\Big)\\
&= \frac{\Gamma(2\kappa-2)2(\pi N_{T_k})^{\kappa-1}}{(N_{T_k}p\tilde{C}r^{\star}_{\mathsf{af}})^{2\kappa-2}}\Big(N_{T_k}\Big(\frac{\big(\alpha_k^{\star}(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1)}dx^1\Big)\\
&= O(1)\frac{N_{T_k}\Big(\frac{\left(\alpha_k^{\star}(N_{T_k})\right)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^{\star}(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1)}dx^1}{(\psi_k(N_{T_k}))^{\kappa-1}} = o(1),
\end{aligned}
$$

by application of Lemma 6 and where $\Gamma(\kappa)=\int_0^{\infty}x^{\kappa-1}\mathrm{e}^{-x}dx$ denoting the gamma function.

Step 2. We show now that the same result holds for $(x^1,\ldots,x^{\kappa})\in\mathcal{C}_{\kappa\mathbf{n}}$ for any $2\leq\mathbf{n}<\kappa$. For any random $\mathbf{n}-$partition

$$\Pi_{\kappa}=\{P_1,\ldots,P_{\mathbf{n}}\}\qquad\text{of the subset }\ [\kappa]:=\{1,\ldots,\kappa\},$$

where each component P_j, $j=1,\ldots,\mathbf{n}$, is of cardinal $|P_j|$ and let denote by $\bar{\mathcal{D}}^{\kappa}(\Pi_{\kappa})$ the set of $(x^1,\ldots,x^{\kappa})\in(\bar{\mathscr{D}}_G)^{\kappa}$ such that the vertices $\{x^i:i\in P_j\}$ formed a connected component of $\mathbb{S}_{2r^{\star}_{\mathsf{af}}}(x^1,\ldots,x^{\kappa})$. Hence,

$$\mathcal{C}_{\kappa\mathbf{n}}=\bigcup_{\text{all }\mathbf{n}-\text{partitions }\Pi_{\kappa}}\bar{\mathcal{D}}^{\kappa}(\Pi_{\kappa}),$$

and it suffice to prove that for any $\mathbf{n}-$partition Π_κ

$$(N_{T_k})^\kappa \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big) \int_{\bar{\mathcal{D}}^\kappa(\Pi_\kappa)} \mathrm{e}^{-N_{T_k} p\mathrm{vol}(x^1,\ldots,x^\kappa)} dx^1 \cdots dx^\kappa \longrightarrow 0,$$

as both $\alpha, k \to \infty$.

Second, without loss of generality, let now fix one arbitrary $\mathbf{n}-$partition Π_κ and observe that for any $(x^1,\ldots,x^\kappa) \in \bar{\mathcal{D}}^\kappa(\Pi_\kappa)$, we have

$$\bar{\mathcal{D}}^\kappa(\Pi_\kappa) \subseteq \prod_{j=1}^{\mathbf{n}} \mathcal{C}_{|P_j|1} \qquad \text{and} \qquad \mathrm{vol}(x^1,\ldots,x^\kappa) = \sum_{j=1}^{\mathbf{n}} \mathrm{vol}(\{x^i : i \in P_j\}). \tag{5.7}$$

It follows from (5.7) that

$$\begin{aligned}
&(N_{T_k})^\kappa \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big) \int_{\bar{\mathcal{D}}^\kappa(\Pi_\kappa)} \mathrm{e}^{-N_{T_k} p\mathrm{vol}(x^1,\ldots,x^\kappa)} dx^1 \cdots dx^\kappa \\
&\quad = \Big(N_{T_k} \frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big)^\kappa \int_{\bar{\mathcal{D}}^\kappa(\Pi_\kappa)} \prod_{j=1}^{\mathbf{n}} \mathrm{e}^{-N_{T_k} p\mathrm{vol}(\{x^i : i\in P_j\})} dx^1 \cdots dx^\kappa \\
&\quad \le \Big(N_{T_k} \frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big)^\kappa \prod_{j=1}^{\mathbf{n}} \int_{\mathcal{C}_{|P_j|1}} \mathrm{e}^{-N_{T_k} p\mathrm{vol}(\{x^i : i\in P_j\})} \prod_{i\in P_j} dx^i \\
&\quad = \prod_{j=1}^{\mathbf{n}} \Big(N_{T_k} \frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big)^{|P_j|} \int_{\mathcal{C}_{|P_j|1}} \mathrm{e}^{-N_{T_k} p\mathrm{vol}(\{x^i : i\in P_j\})} \prod_{i\in P_j} dx^i,
\end{aligned}$$

which tends to zero as shown in Step 1, completing the proof. □

Next we study the limit in $\mathcal{C}_{\kappa\kappa}$.

Lemma 8 *Assume that A2 and A3 are satisfied. If $r_{\mathsf{af}}^\star(N_{T_k})$ is given by (3.14) and $\alpha_k^\star$ is given by (3.8) and satisfying (3.11), then, for any $\kappa \ge 2$ and $(x^1,\ldots,x^\kappa) \in \mathcal{C}_{\kappa\kappa}$, we have*

$$(N_{T_k})^\kappa \Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!} \mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)} \Big) \int_{\mathcal{C}_{\kappa\kappa}} \mathrm{e}^{-N_{T_k} p\mathrm{vol}(x^1,\ldots,x^\kappa)} dx^1 \cdots dx^\kappa \overset{\alpha,k\to+\infty}{\longrightarrow} \mathrm{e}^{-\kappa \lim\limits_{\alpha,k\to\infty} \mathfrak{f}_k(N_{T_k})}.$$

Proof Recall that for any $(x^1,\ldots,x^\kappa)\in\mathcal{C}_{\kappa\kappa}$ we observe that $\mathrm{vol}(x^1,\ldots,x^\kappa)=\sum_{i=1}^{\kappa}\mathrm{vol}(x^i)$. This enables us to write

$$\begin{aligned}
&(N_{T_k})^\kappa\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{\mathcal{C}_{\kappa\kappa}}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1,\ldots,x^\kappa)}dx^1\cdots dx^\kappa\\
&\quad=(N_{T_k})^\kappa\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\Big\{\int_{(\bar{\mathscr{D}}_G)^\kappa}\mathrm{e}^{-N_{T_k}p\sum_{i=1}^{\kappa}\mathrm{vol}(x^i)}dx^1\cdots dx^\kappa\\
&\qquad-\int_{(\bar{\mathscr{D}}_G)^\kappa\setminus\mathcal{C}_{\kappa\kappa}}\mathrm{e}^{-N_{T_k}p\sum_{i=1}^{\kappa}\mathrm{vol}(x^i)}dx^1\cdots dx^\kappa\Big\}\\
&\quad=\Big\{\prod_{i=1}^{\kappa}\Big(N_{T_k}\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-1}}{(N_{T_k}-1)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{\bar{\mathscr{D}}_G}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^i)}dx^i\Big\}\\
&\qquad-(N_{T_k})^\kappa\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{(\bar{\mathscr{D}}_G)^\kappa\setminus\mathcal{C}_{\kappa\kappa}}\mathrm{e}^{-N_{T_k}p\sum_{i=1}^{\kappa}\mathrm{vol}(x^i)}dx^1\cdots dx^\kappa \qquad (5.8)\\
&\quad\overset{\alpha,k\to+\infty}{\longrightarrow}\Big(\mathrm{e}^{-\lim\limits_{\alpha,k\to\infty}\mathfrak{f}_k(N_{T_k})}\Big)^\kappa-0.
\end{aligned}$$

There may be some doubt as to why the term (5.8) tends to zero. Let us check this by observing that for any $(x^1,\ldots,x^\kappa)\in(\bar{\mathscr{D}}_G)^\kappa\setminus\mathcal{C}_{\kappa\kappa}$ we have $\mathrm{vol}(x^1,\ldots,x^\kappa)\leq\sum_{i=1}^{\kappa}\mathrm{vol}(x^i)$ which enables us to find

$$\begin{aligned}
&(N_{T_k})^\kappa\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{(\bar{\mathscr{D}}_G)^\kappa\setminus\mathcal{C}_{\kappa\kappa}}\mathrm{e}^{-N_{T_k}p\sum_{i=1}^{\kappa}\mathrm{vol}(x^i)}dx^1\cdots dx^\kappa\\
&\quad\leq(N_{T_k})^\kappa\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{(\bar{\mathscr{D}}_G)^\kappa\setminus\mathcal{C}_{\kappa\kappa}}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1,\ldots,x^\kappa)}dx^1\cdots dx^\kappa\\
&\quad=\sum_{\mathbf{n}=1}^{\kappa-1}\Big\{(N_{T_k})^\kappa\Big(\frac{\big(\alpha_k^\star(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}\mathrm{e}^{-\left(\alpha_k^\star(N_{T_k})\right)}\Big)\int_{\mathcal{C}_{\kappa\mathbf{n}}}\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1,\ldots,x^\kappa)}dx^1\cdots dx^\kappa\Big\},
\end{aligned}$$

where it's straightforward that the sum tends to zero thanks to Lemma 7. □

A key result needed in the sequel for the proofs of Theorems 2 and 3 is as follows:

Lemma 9 *Given a sequence $E_{1k},E_{2k},\ldots,E_{N_{T_k}k}$ of events such that E_{ik} be the event that the vertex $x^i_{T_k}$ is isolated, define $U_{N_{T_k}}$ to be the random number of E_{ik} that hold. If for any set $\{i_1,\ldots,i_\kappa\}$ it is true that*

$$\mathbb{P}\Big(\bigcap_{j=1}^{\kappa}E_{jk}\Big)=\mathbb{P}\Big(\bigcap_{j=1}^{\kappa}E_{i_jk}\Big), \qquad (5.9)$$

and there is a constant u such that for any fixed κ

$$(N_{T_k})^\kappa\times\mathbb{P}\Big(\bigcap_{j=1}^{\kappa}E_{jk}\Big)\overset{\alpha,k\to+\infty}{\longrightarrow}u^\kappa, \qquad (5.10)$$

hence the sequence $U_{N_{T_k}}$ converges in distribution to a Poisson random variable with mean u.

We do not claim originality of the Lemma 9 and, in fact, a similar result has been proved in Alon & Spencer (2000) using a probabilistic version of Brun's sieve theorem and the Bonferroni inequalities. Since we have found this particular result in the literature, we not provide detailed proof.

We now proceed to conclude the proof of Theorem 2.

Proof For $(x^1_{T_k},\ldots,x^\kappa_{T_k})\in\mathcal{C}_{\kappa\kappa}$, we proved in Proposition 3 that

$$\mathbb{P}(E_{1k}\cap\cdots\cap E_{\kappa k}) = \Big(\frac{\big(\alpha^\star_k(N_{T_k})\big)^{N_{T_k}-\kappa}}{(N_{T_k}-\kappa)!}e^{-\left(\alpha^\star_k(N_{T_k})\right)}\Big) \times\int_{\mathcal{C}_{\kappa\kappa}}\Big(1-p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})\Big)^{N_{T_k}-\kappa}dx^1_{T_k}\cdots dx^\kappa_{T_k}.$$

Or, for N_{T_k} almost surely sufficiently large we find

$$\begin{aligned}\Big(1-p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})\Big)^{N_{T_k}-\kappa} &= \frac{\Big(1-p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})\Big)^{N_{T_k}}}{\Big(1-p\sum_{i=1}^\kappa\mathrm{vol}(x^i_{T_k})\Big)^{\kappa}}\\ &= \frac{\Big(1-p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})\Big)^{N_{T_k}}}{\Big(1-p\kappa\pi(r^\star_{\mathsf{af}})^2)\Big)^{\kappa}}\\ &\equiv\frac{\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})}}{1}.\end{aligned}$$

Thus, by application of Lemma 8,

$$(N_{T_k})^\kappa\times\mathbb{P}(E_{1k}\cap\cdots\cap E_{\kappa k})^{\alpha,k\to+\infty}\longrightarrow\Big(\mathrm{e}^{-\lim\limits_{\alpha,k\to\infty}\mathfrak{f}_k(N_{T_k})}\Big)^\kappa=(\phi)^\kappa<\infty,$$

and immediately the condition (5.10) in Lemma 9 is verified. It is easily seen that the condition (5.9) is also verified (its proof is left to the reader) and hence Theorem 2 follows for $(x^1_{T_k},\ldots,x^\kappa_{T_k})\in\mathcal{C}_{\kappa\kappa}$ and generally for $(x^1_{T_k},\ldots,x^\kappa_{T_k})\in(\bar{\mathscr{D}}_G)^\kappa$ since for $(x^1_{T_k},\ldots,x^\kappa_{T_k})\in(\bar{\mathscr{D}}_G)^\kappa\setminus\mathcal{C}_{\kappa\kappa}$ and for N_{T_k} almost surely sufficiently large

$$\Big(1-p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})\Big)^{N_{T_k}-\kappa}\leq\frac{\Big(1-p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})\Big)^{N_{T_k}}}{\Big(1-p\kappa\pi(r^\star_{\mathsf{af}})^2)\Big)^{\kappa}}\equiv\mathrm{e}^{-N_{T_k}p\mathrm{vol}(x^1_{T_k},\ldots,x^\kappa_{T_k})}.$$

Then, the probability of the event $(E_{1k}\cap E_{2k}\cap\cdots\cap E_{\kappa k})$ times $(N_{T_k})^\kappa$ tends to zero as $\alpha,k\to\infty$ by Lemma 7, this completes the proof. □

Finally, we conclude the proof of Theorem 3.

Proof Notice that

$$\mathbb{P}(F_{1k}\cap\cdots\cap F_{\kappa k})=p^\kappa\mathbb{P}(E_{1k}\cap\cdots\cap E_{\kappa k}),$$

and immediately Theorem 3 follows using the same arguments as for Theorem 2. □

6 Discussion

In this framework, a measure-valued process is constructed for modeling a dynamic social network on geometric graph. The distribution of isolated vertices (active and inactive) is investigated by a thorough analysis and numerical experiments are proposed. We presume a social network is composed of particles in interaction in continuous time represented by a set of nodes over some state space (an open of $\mathbb{R}^d$). We assumed a Bernoulli probability $0 < p \leq 1$ on the activity state of each node in the network. We proved that if the state space is the unit square and all nodes have the same radius of affinity $r^{\star}_{\mathsf{af}} = \sqrt{\frac{(N_t)^l}{p\pi N_t}}$ for $l \in (0,1)$, the number of isolated nodes is asymptotically Poisson with mean ϕ where $\phi \sim 0$ (at time T_k) as the current size $N_{T_k} \in \mathbb{N}$ of the network tends almost surely to infinity (here when the invitation rate of particles and the time interval are both large; $\alpha, k \to \infty$). This offers a natural threshold for the construction of a detection procedure of isolated nodes tailored to the dynamic clustering of the network adaptively from the data. The question of whether vanishment of isolated nodes almost surely ensures connectivity of the network or not remains open.

Availability of Data and Acknowlegement

The data that support the findings of this study are available on request from the corresponding author, [K. Khadraoui; khader.khadraoui@mat.ulaval.ca]. A. Sid-Ali would like to thank the Institut des Sciences Mathématiques for partially supporting his PhD. This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

Conflict of interest statement and Author contribution

All authors declare that they have no conflicts of interest. All authors wrote, reviewed and edited the manuscript.

References

Aldous D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, **25**, 812–854.

Alon N. & Spencer J.H. (2000). *The probabilistic method.* 2nd edn. Wiley.

Arias-Castro E. & Verzelen N. (2014). Community detection in random networks. *Annals of Statistics*, **42**, 940–969.

Atar R., Kang W., Kaspi H. & Ramanan K. (2021). Large-time limit of nonlinearly coupled measure-valued equations that model many-server queues with reneging. *Preprint, https://arxiv.org/abs/2107.05226.*

Bickel P.J. & Chen A. (2009). A nonparametric view of network models and newman-girvan and other modularities. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 21068–21073.

Bickel P.J., Chen A., Zhao Y., Levina E. & Zhu J. (2015). Correction to the proof of consistency of community detection. *The Annals of Statistics*, **43**, 462–466.

Bollobás B. (2001). *Random graphs.* 2nd edn. Cambridge Univ. Press.

Bollobás B., Janson S. & Riordan O. (2007). The phase transition in inhomogeneous random graphs. *Rand. Struct. Algs*, **31**, 3–122.

Borgs C., Chayes J.T., Cohn H. & Holden N. (2018). Sparse exchangeable graphs and their limits via graphon processes. *Journal of Machine Learning Research*, **18**, 1–71.

Boucheron S., Lugosi G. & Massart P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence.* Univ. Press, Oxford.

Caron F. & Fox E.B. (2017). Sparse graphs using exchangeable random measures. *Journal of the Royal Statistical Society: Series B*, **79**, 1295–1366.

Carvalho H., Del Moral P., Monin A. & Salut G. (1997). Optimal nonlinear filtering in gps/ins integration. *IEEE Transactions on Aerospace and Electronic Systems*, **33**(3), 835–850.

Chatterjee S. & Durrett R. (2009). Contact processes on random graphs with degree power law distribution have critical value zero. *Ann. Probab.*, **37**, 2332–2356.

Crisan D. & Lyons T. (1997). Nonlinear filtering and measure-valued processes. *Probability Theory and Related Fields*, **109**(2), 217–244.

Dawson D. (1993). Measure-valued markov processes. Dans *École d'été de Probabilités de Saint-Flour XXI-1991. Lecture Notes in Mathematics, Vol. 1541*, réd. P.H. (eds), pp. 1–260. Springer-Verlag, Berlin, Heidelberg.

Dawson D.A. (2017). *Introductory Lectures on Stochastic Population Systems.* Preprint, ArXiv:1705.03781 [math.PR].

Decreusefond L., Dhersin J.S., Moyal P. & Tran V.C. (2012). Large graph limit for an sir process in random network with heterogeneous connectivity. *Ann. Appl. Probab.*, **22**(2), 541–575.

Del Moral P. (1996). Nonlinear filtering: interacting particle resolution. *Markov Processes and Related Fields*, **2**(4), 555–580.

Del Moral P. & Doucet A. (2014). Particle methods: An introduction with applications. *ESAIM: Proceedings.*, **44**, 1–46.

Dette H. & Henze N. (1989). The limit distribution of the largest nearest neighbor link in the unit d-cube. *Journal of Applied Probability*, **26**, 67–80.

Dobrushin R.L. (1971). Markov processes with a large number of locally interacting components: existence of a limit process and its ergodicity. *Problems of Information Transmission*, **7**(2), 149–164.

Durrett R. (2007). *Random Graph Dynamics.* Cambridge Univ. Press, Cambridge.

Finkelshtein D., Kondratiev Y. & Kutoviy O. (2009). Individual based model with competition in spatial ecology. *SIAM: Journal on Mathematical Analysis*, **41**, 297–317.

Fournier N. & Méléard S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Annals of Applied Probability*, **14**, 1880–1919.

Gilbert E.N. (1961). Random plane networks. *J. Soc. Indust. Appl. Math.*, **9**, 533–543.

Holland P.W., Laskey K.B. & Leinhardt S. (1983). Stochastic blockmodels: First steps. *Social Networks*, **5**, 109–137.

Jin J. (2015). Fast community detection by score. *The Annals of Statistics*, **43**, 57–89.

Kang W. & Ramanan K. (2010). Fluid limits of many-server queues with reneging. *Annals of Applied Probability*, **20**, 2204–2260.

Khadraoui K. & Sid-Ali A. (2022). Weak convergence of a measure-valued process for social networks. *Preprint, arXiv:1809.04388*.

Kisdi E. (1999). Evolutionary branching under asymmetric competition. *J. Theor. Biol.*, **197**.

Liggett T. (1985). *Interacting Particle Systems*. Springer, New York.

McKean H.P. (1967). Propagation of chaos for a class of non-linear parabolic equations. Dans *Lecture Series in Differential Equations*, vol. 7, pp. 41–57. Catholic Univ.

Méléard S. & Roelly-Coppolett S. (1987). A propagation of chaos result for a system of particles with moderate interaction. *Stochastic Process. Appl.*, **26**, 317–332.

Mountford T., Valesin D. & Yao Q. (2013). Metastable densities for contact processes on random graphs. *Electron. J. Probab.*, **103**, 1–36.

Norros I. & Reittu H. (2006). On a conditionally Poissonian graph process. *Adv. Appl. Probab.*, **38**, 59–75.

Park Y. & Bader J. (2012). How networks change with time. *Bioinformatics*, **28**, i40–48.

Penrose M.D. (1997). The longest edge of the random minimal spanning tree. *The Annals of Applied Probability*, **7**, 340–361.

Penrose M.D. (2003). *Random Geometric Graphs*. Univ. Press, Oxford.

Penrose M.D. (2016). Connectivity of soft random geometric graphs. *The Annals of Applied Probability*, **26**, 986–1028.

Perry P.O. & Wolfe P.J. (2013). Point process modelling for directed interaction networks. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, **75**, 821–849.

Schapira B. & Valesin D. (2017). Extinction time for the contact process on general graphs. *Probab. Theory Related Fields*, **3**, 871–899.

Snijders T.A.B. & Nowicki K. (1997). Estimation and prediction for stochastic block-models for graphs with latent block structure. *Journal of classification*, **14**, 75–100.

Spitzer F. (1969). Random processes defined through the interaction of an infinite particle system. Dans *Lecture Notes in Math*, vol. 89, pp. 201–223. Springer, Berlin.

Sznitman A.S. (1984). Nonlinear reflecting diffusion process and the propagation of chaos and fluctuations associated. *J. Functional Anal.*, **56**, 311–336.

Tran V.C. (2008). Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM: Probability and Statistics*, **12**, 345–386.

van der Hofstad R. (2017). *Random Graphs and Complex Networks*. Cambridge Univ. Press, Cambridge.

Veitch V. & Roy D.M. (2015). The class of random graphs arising from exchangeable random measures. *Preprint arXiv: 1512.03099*.

Yuan B., Li H., Bertozzi A.L., Brantingham P.J. & Porter M.A. (2019). Multivariate spatiotemporal hawkes processes and network reconstruction. *SIAM Journal on Mathematics of Data Science*, **01**, 356–382.

Zhao Y., Levina E. & Zhu J. (2012). Consistency of community detection in networks under degree-corrected stochastic block models. *The Annals of Statistics*, **40**, 2266–2292.