

Large deviations for the greedy exploration process on configuration models

Bermolen Paola* Goicoechea Valeria† Jonckheere Matthieu‡

Abstract

We prove a large deviation principle for the greedy exploration of configuration models, building on a time-discretized version of the method proposed by [2] and [4] for jointly constructing a random graph from a given degree sequence and its exploration. The proof of this result follows the general strategy to study large deviations of processes proposed by [9], based on the convergence of non-linear semigroups.

We provide an intuitive interpretation of the LD cost function using Crámer's theorem for the average of random variables with appropriate distribution, depending on the degree distribution of explored nodes.

The rate function can be expressed in a closed-form formula, and the large deviations trajectories can be obtained through explicit associated optimization problems. We then deduce large deviations results for the size of the independent set constructed by the algorithm. As a particular case, we analyze these results for d -regular graphs.

Keywords: large deviations; random graphs; configuration model; comparison principle; Hamilton-Jacobi equations.

MSC2020 subject classifications: 60.

Submitted to ECP on December 23, 2021, final version accepted on August 19, 2023.

1 Introduction

In this paper, we analyze a simple greedy algorithm to construct an independent set over a random graph chosen uniformly from those with a given degree. We use a simultaneous construction of the random graph from a given degree sequence (i.e., it is a *configuration model*) and an exploration discovering independent set. This idea was first used by [15] for d -regular graphs and then for [2] and [4] for more general uniform random graphs. We consider a time-discretized version of the algorithm proposed by [4] for a bounded degree sequence. We use this discrete-time version of the process because it simplifies the analysis to obtain the large deviations for the size of the independent set obtained.

We prove a large deviation principle (LDP), when the number N of the graph's vertices goes to infinity, for a rescaled version $X_t^N = \frac{Y_{[Nt]}^N}{N}$ ($t \in [0, 1]$) of the multidimensional Markov chain $\{Y_n^N\}_n$ that counts the number of vertices that have already been assigned

*Universidad de la República, Uruguay. E-mail: paola@fing.edu.uy

†Universidad de la República, Uruguay. E-mail: vgoicoechea@fing.edu.uy

‡LAAS CNRS, France. E-mail: matthieu.jonckheere@laas.fr

into the independent set, and the number of *empty* (or non-explored) vertices from each degree at each step n of the algorithm.

The proof of this result is based on the general strategy proposed by Feng and Kurtz in [9], which is based on the convergence of non-linear semigroups. Their seminal work consists of combining the tools of probability, analysis, and control theory used in the works of [6, 7, 8, 10, 13], and others to propose a general strategy for the study of large deviations of processes. In the case of Markov processes, this program is carried out in four steps. The first step consists of proving the convergence of non-linear generators \mathbf{H}_N and derives the limit operator \mathbf{H} . The second step consists of verifying the *exponential compact containment condition*. The third step consists of proving that \mathbf{H} generates a semigroup $\mathbf{V} = \{V_t\}_t$. The rate function is constructed in terms of that limit \mathbf{V} . This limit semigroup usually admits a variational form known as *Nisio semigroup* in control theory. Then, the fourth step consists in constructing a version of the rate I that is much more tractable in practice.

In our case, after working on the four steps mentioned before, we prove that the rate function can be expressed as an action integral. Moreover, its cost function has an intuitive interpretation in terms of Crámer's theorem for the average of random variables with appropriate distribution, depending on the degree distribution of explored vertices in each step of the algorithm.

We provide a way to find the trajectory that minimizes the LD rate function over a set of trajectories (i.e., the most probable trajectory) by studying the Hamiltonian dynamics associated with the rate function obtained and deduce LD results for the size of the independent set constructed by such an algorithm.

The rest of the paper is organized as follows. In Sect. 2, we define the dynamic analyzed in this article, which consists of simultaneously constructing a random graph and an independent set from an initial distribution of degrees. Moreover, we define a sequence of Markov processes related to this algorithm. In Sect. 3, we present the main result of this article: a path-state LDP for the sequence of Markov processes defined in Sect. 2 along with the heuristic that motivates the result. An outline of the proof is deferred to Sect. 5. As a corollary, we obtain the fluid limit of the process and LD results for the size of the independent set constructed by the algorithm. In Sect. 4, we analyze those results for the particular case of d -regular graphs.

2 Description of the dynamics

This section presents the dynamics considered in this article, which consist of a simultaneous construction of a random graph and an independent set from an initial distribution of degrees.

We start with a set of vertices $\mathcal{V}^N = \{1, 2, \dots, N\}$ where each vertex i has a degree $\deg(i)$ assigned such that $\deg(i) \leq D < \infty$ for all i and such that the initial distribution of degrees $\frac{1}{N} \# \{i : \deg(i) = j\}$ converges to $p_j \geq 0$, when the number of vertices N goes to infinity, for all $j = 0, \dots, D$ with $\sum_{j=0}^D p_j = 1$. For the existence and construction of such graphs, see for instance [14]. Let us denote $\lambda = \sum_j j p_j$.

Each vertex i of the graph has a number $\deg(i)$ of half-edges available to be paired with the half-edges of other vertices. Next we describe how these half-edges are paired as the random graph is constructed.

At each step $n = 0, 1, \dots, N$, the set \mathcal{V}^N is partitioned into three classes:

- a set \mathcal{S}_n^N of vertices that have already been assigned into the independent set, with all half-edges paired with vertices out of \mathcal{S}_n^N ;
- a set \mathcal{B}_n^N of *blocked vertices*, where at least one of its half-edges has been paired with a half-edge from \mathcal{S}_n^N ;

- a set \mathcal{E}_n^N of *empty vertices*, from which no half-edge has yet been paired. \mathcal{E}_n^N can be decomposed as $\mathcal{E}_n^N = \bigcup_{j=0}^D \mathcal{E}_n^N(j)$, where $\mathcal{E}_n^N(j)$ is the set of empty vertices of degree j at step n .

Initially, all vertices are *empty*, i.e. $\mathcal{E}_0^N = \mathcal{V}^N$ and $\mathcal{S}_0^N = \emptyset$. At step n , a vertex v is selected uniformly from \mathcal{E}_n^N , it is assigned to \mathcal{S}_n^N , and all its half-edges are sequentially paired, drawing uniformly within the available half-edges. This pairing procedure results in the following updates:

- v is moved from \mathcal{E}_n^N to \mathcal{S}_n^N ,
- each half-edge incident to v (if it has some) is paired in turn with some other uniformly randomly chosen vertices among the currently unpaired half-edges,
- all vertices in \mathcal{E}_n^N with some half-edges already paired with a half-edge from v are moved to \mathcal{B}_n^N .

Note that some half-edges from v may be paired with half-edges from \mathcal{B}_n^N , or indeed with other half-edges from v , and no change in the status of a vertex results from such pairings. At each step n , the only paired edges are those with at least one endpoint in \mathcal{S}_n^N . This is the main difference between the dynamics described in [2] and [4] for a continuous-time version of this algorithm. In [2], the neighbors of blocked vertices are revealed at each step, meaning that the number of half-edges available to be paired from an empty vertex may not agree with its initial degree (and changes over time). For simplicity, we do not do this here.

The algorithm terminates at the first step $n = T_N^*$ at which $\mathcal{E}_n^N = \emptyset$. At this point, there may still be some unpaired half-edges pointing out from blocked vertices. These may be paired off uniformly at random to complete the construction of the graph $G(N, (\deg(1), \dots, \deg(N)))$. Note that T_N^* coincides with the size of the independent set constructed by the algorithm. The expected value of $\frac{T_N^*}{N}$ is usually called the *jamming constant* of the graph.

For each $n \in \{0, 1, \dots, N\}$, let us denote $Y_n^N = (S_n^N, U_n^N, E_n^N(0), E_n^N(1), \dots, E_n^N(D))$, being

- $S_n^N = |\mathcal{S}_n^N|$, the number of vertices that have already been assigned into the independent set at step n ;
- U_n^N , the total number of unpaired half-edges (corresponding to empty or blocked vertices) at step n ;
- $E_n^N(j) = |\mathcal{E}_n^N(j)|$, the number of empty vertices with degree j at step n .

$\{Y_n^N\}_n$ is a discrete-time Markov process in \mathbb{R}^{D+3} . By construction, it is updated at step $n + 1$ as follows:

- The vertex v is assigned to \mathcal{S}_n^N . Then, $S_{n+1}^N = S_n^N + 1$.
- If $v \in \mathcal{E}_n^N(k)$ with $k \neq 0$, then:
 1. Each one of the k half-edges pointing out from v is paired in turn with some other currently unpaired half-edges chosen uniformly at random. We are only interested in the number of half-edges from v paired with another vertex different from v (blocked or empty), i.e., that do not form loops. Then, we can assume that we have a pool consisting of k black arrows representing the k half-edges pointing out from v and $U_n^N - k$ white arrows representing the other half-edges available at step n . From this pool, we take sequentially and without replacement k arrows, representing the k half-edges to be paired with the k half-edges of v . Let H^N be the number of selected white half-edges. Then,

H^N has a hypergeometric distribution $\mathbf{Hyper}(U_n^N, U_n^N - k, k)$. Note, however, that $k - H^N$ could be odd (i.e., not all the remaining $k - H^N$ half-edges could form a loop). The solution, in this case, is to pair $k - H^N - 1$ edges to form loops (we do not care which ones), and the remaining half-edge is paired with a new white arrow taken from the pool. This means that the number of white arrows taken from the pool will eventually be $H^N + 1$. For this reason, we define a new random variable \tilde{H}^N such that $\tilde{H}^N = \begin{cases} H^N, & \text{if } k - H^N \text{ is even,} \\ H^N + 1, & \text{if } k - H^N \text{ is odd.} \end{cases}$

Then, $U_{n+1}^N = U_n^N - k - \tilde{H}^N$.

2. We have to distribute those \tilde{H}^N half-edges between the $U_n^N - \sum_j jE_n^N(j)$ half-edges corresponding to blocked vertices and the $\sum_j jE_n^N(j)$ half-edges corresponding to empties. Let B^N be the number of half-edges of v that are paired to blocked vertices, then B^N has a hypergeometric distribution $\mathbf{Hyper}(U_n^N - k, U_n^N - \sum_{j=1}^D jE_n^N(j), \tilde{H}^N)$.

3. Now, if $\tilde{H}^N = h$ (with $h \leq k$) and $B^N = b$ (with $b \leq h$), there are $h - b$ half-edges to distribute among the empties. Let W_j^N be the number of half-edges from v that are paired to some $w \in \mathcal{E}_n^N(j)$. Note that (W_1^N, \dots, W_D^N) has a (multivariate) hypergeometric distribution:

$$\mathbf{Hyper}\left(\sum_j jE_n^N(j) - k, E_n^N(1), \dots, k(E_n^N(k) - 1), \dots, DE_n^N(D), h - b\right).$$

4. Finally, let \tilde{W}_j^N be the number of empty vertices of degree j that share at least one edge with v . Then, $E_{n+1}^N(0) = E_n^N(0)$, $E_{n+1}^N(j) = E_n^N(j) - \tilde{W}_j^N$ if $j \neq k$ and $E_{n+1}^N(k) = E_n^N(k) - 1 - \tilde{W}_k^N$.

- If $\mathbf{deg}(v) = 0$, then $S_{n+1}^N = S_n^N + 1$, $U_{n+1}^N = U_n^N$, $E_{n+1}^N(0) = E_n^N(0) - 1$ and $E_{n+1}^N(j) = E_n^N(j)$ for all $j \neq 0$.

According to the following Lemma, the distribution of $(\tilde{W}_1, \dots, \tilde{W}_D)$ can be approximated by the hypergeometric distribution corresponding to (W_1, \dots, W_D) . This simplifies the calculation of limits related to the linear generator T^N of the Markov chain $\{Y_n^N\}_n$, which we use as part of the strategy for studying the large deviations for a rescaled version of $\{Y_n^N\}_n$.

Lemma 2.1. *Let $x = (\hat{s}, \hat{u}, \hat{e}_0, \dots, \hat{e}_D)$ be an element in the space state of $\{Y_n^N\}_n$, and $(\omega_j)_j$ with $0 \leq \omega_j \leq \hat{e}_j$ such that $\sum_j \omega_j \leq h - b$. Then,*

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(\left(\tilde{W}_j^N\right)_j = (\omega_j)_j \mid Y_n^N = x; \mathbf{deg}(v) = k; \tilde{H}^N = h; B^N = b; (W_j^N)_j = (\omega_j)_j\right) = 1$$

Proof. See Equation 17 from [2]. In the article notation: $W_j^N = Y(\mu_{t^-})(j)$ and $\tilde{W}_j^N = \tilde{Y}(\mu_{t^-})(j)$. □

Let $X_t^N := \frac{Y_{[Nt]}}{N}$ be a rescaled version of Y_n^N with $t \in [0, 1]$. The state space of X_t^N is $M^N = \left\{ \frac{1}{N}(\hat{s}, \hat{u}, \hat{e}_0, \dots, \hat{e}_D) : \hat{s}, \hat{e}_i \in \{0, \dots, N\}, \hat{u} \in \mathbb{N}, \sum_j j\hat{e}_j \leq \hat{u} \leq \lambda N \right\}$ which is a subset of the compact set $M := \left\{ (s, u, e_0, e_1, \dots, e_D) \in [0, 1] \times \mathbb{R} \times [0, 1]^{D+1} : \sum_j je_j \leq u \leq \lambda \right\}$. Note that $X^N \in D_M[0, 1]$, the space of càdlàg functions $\mathbf{x} : [0, 1] \rightarrow M$. We consider this space equipped with the Skorohod topology. The size of the independent set constructed by such an algorithm is given by $T_N^* = \inf \left\{ n : \sum_j E_n^N(j) = 0 \right\}$. We provide large deviations for both sequences $\{X_t^N\}_N$ and $\left\{ \frac{T_N^*}{N} \right\}_N$.

3 Main results

In this section, we present the main results of the paper. In Sect. 3.1, we present an LDP for $X^N = \{X_t^N\}_{0 \leq t \leq 1}$ and a heuristic description to derive this result. The proof of this theorem is based on the work of [9] and is deferred to Sect. 5. In Sect. 3.2, we deduce the corresponding fluid limit. In Sect. 3.3, we provide a way of finding the trajectory that minimizes the LD rate function over a set of trajectories (i.e., the most probable trajectory) by studying the Hamiltonian dynamics associated with the rate function obtained. Finally, in Sect. 3.4, we deduce an LD result for the size of the independent set constructed by such an algorithm.

3.1 LDP for $\{X^N\}_N$

We now state our main result. Let (p_0, p_1, \dots, p_D) be a probability distribution and let $\lambda = \sum_{k=0}^D k p_k$. Let $\mathcal{V}^N = \{1, 2, \dots, N\}$ be a given set of vertices with an assignation of degrees such that $\deg(i) \leq D < \infty$ for all i and such that $\frac{1}{N} \# \{i : \deg(i) = j\} \rightarrow p_j$ when N goes to infinity.¹ Consider the algorithm defined in Sect. 2 to sequentially construct the graph from such a degree configuration while constructing the independent set.

Let $Y_n^N = (S_n^N, U_n^N, E_n^N(0), E_n^N(1), \dots, E_n^N(D))$ be the Markov chain that keeps track of the number of vertices in the independent set, the number of available half-edges and the number of empty vertices of each degree at each step n of the algorithm. Let $X_t^N := \frac{Y_{\lfloor Nt \rfloor}^N}{N}$ be a rescaled version of Y_n^N with $t \in [0, 1]$.

Theorem 3.1 (LDP for $\{X^N\}_N$). *The sequence $\{X^N\}_N$ with $X^N = \{X_t^N\}_{0 \leq t \leq 1}$ verifies an LDP on $D_M[0, 1]$ with good rate function $I: D_M[0, 1] \rightarrow [0, +\infty]$ such that $I(\mathbf{x}) = \int_0^1 L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$ if $\mathbf{x} \in \mathcal{H}_L$ and it is $+\infty$ in other case. $L: M \times \mathbb{R}^{D+3} \rightarrow \mathbb{R}$ is the cost function*

$$L(x, \beta) = \sup_{\alpha \in \mathbb{R}^{D+3}} \{ \langle \alpha, \beta \rangle - H(x, \alpha) \}, \tag{3.1}$$

with $H: M \times \mathbb{R}^{D+3} \rightarrow \mathbb{R}$ given by

$$H(x, \alpha) = \begin{cases} \log \left[\sum_k e^{\alpha_s - 2k\alpha_u - \alpha_k} \left(1 + \sum_{j=1}^D (e^{-\alpha_j} - 1) \frac{j e_j}{u} \right)^k \frac{e_k}{\sum_j e_j} \right], & \text{if } \sum_j e_j > 0, \\ 0, & \text{if } \sum_j e_j = 0, \end{cases} \tag{3.2}$$

being $x = (s, u, e_0, e_1, \dots, e_D)$ and $\alpha = (\alpha_s, \alpha_u, \alpha_0, \alpha_1, \dots, \alpha_D)$. \mathcal{H}_L is the set of all absolutely continuous function $\mathbf{x}: [0, 1] \rightarrow M$, $\mathbf{x}(t) = (s(t), u(t), e_0(t), e_1(t), \dots, e_D(t))$ with initial value $\mathbf{x}(0) = (0, \lambda, p_0, p_1, \dots, p_D)$ and such that $s(t)$ is increasing, $u(t)$ and $e_j(t)$ are decreasing, and the integral $\int_0^1 L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) dt$ exists and it is finite.

An outline of the proof of this result is deferred to Sect. 5. In what follows, we provide an intuitive interpretation of the cost function of the LDP in terms of the rate function provided by Cramer’s theorem for the average of the approximated distribution of the new explored vertices in one step, conditioning to the number of explored vertices. Then, the following lines are intended to give an informal interpretation of the cost function obtained in Theorem 3.1. For this reason, with a slight abuse of notation, we use the symbol \approx to refer to different types of approximations (e.g. for the definition of the derivative, to say that the process X_t^N is close to the path $\mathbf{x}(t)$ in a fixed time t , and so on).

Consider a curve $\mathbf{x}(t) = (s(t), u(t), e_0(t), \dots, e_D(t))$ contained in M and such that $X_t^N \approx \mathbf{x}(t)$. Then, the infinitesimal increments $\dot{\mathbf{x}}(t)$ correspond to the mean number

¹This is a deterministic convergence and we assume that there exists a graph satisfying this condition.

of new explored vertices from each degree in one step, as can be deduced from the following statement:

$$\dot{\mathbf{x}}(t) \approx \frac{\mathbf{x}(t+h) - \mathbf{x}(t)}{h} \approx \frac{Y_{[N(t+h)]}^N - Y_{[Nt]}^N}{Nh} = \frac{1}{Nh} \sum_{n=[Nt]}^{[N(t+h)]-1} (Y_{n+1}^N - Y_n^N).$$

Proposition 3.2. *The distribution of the number of new explored vertices in one step $Y_{n+1}^N - Y_n^N$, conditioning to $X_t^N \approx x(t) = (s(t), u(t), e_0(t), \dots, e_D(t))$, can be approximated by the random vector:*

$$Z^{\mathbf{x}(t)} = \begin{cases} (1, 0, -1, 0, \dots, 0), & \text{with probability } \frac{e_0(t)}{\sum_j e_j(t)}, \\ (1, -2k, 0, -M_1, \dots, -1 - M_k, \dots, -M_D), & \text{with probability } \frac{e_k(t)}{\sum_j e_j(t)} \quad (1 \leq k \leq D), \end{cases}$$

where $M \sim \text{Mult}(K - B, q_1, \dots, q_D)$ is a multinomial vector depending on K such that $\mathbb{P}(K = k) = \frac{e_k(t)}{\sum_j e_j(t)}$ for $k \in \{0, \dots, D\}$, $B = B(K) \sim \text{Bin}\left(K, 1 - \frac{\sum_j j e_j(t)}{u(t)}\right)$, and $q_i = \frac{i e_i(t)}{\sum_j j e_j(t)}$.

Proof. At step n , the vertex $v \in \mathcal{E}_n^N(k)$ is drawn uniformly. If $k \neq 0$, then $Y_{n+1}^N - Y_n^N = (1, -k - \tilde{H}^N, 0, -\tilde{W}_1^N, \dots, -1 - \tilde{W}_k^N, \dots, -\tilde{W}_D^N)$, where

- \tilde{H}^N is the number of half-edges from v that are joined to another node different from v . As the probability of loops converges to zero (see [4]), then $\tilde{H}^N \approx k$.
- Lemma 2.1 assures that $\tilde{W}_j^N \approx W_j^N$, where (W_1^N, \dots, W_D^N) has a (multivariate) hypergeometric distribution

$$\text{Hyper}\left(\sum_j j E_n^N(j) - k, E_n^N(1), \dots, j E_n^N(j), \dots, k (E_n^N(k) - 1), \dots, D E_n^N(D), k - B^N\right).$$

Note that B^N can be approximated by a binomial random variable

$$B \sim \text{Bin}\left(k, p = \lim_N \frac{U_n^N - \sum_j j E_n^N(j)}{U_n^N - k} = 1 - \frac{\sum_j j e_j(t)}{u(t)}\right).$$

Moreover, (W_1^N, \dots, W_D^N) can be approximated by a multinomial random vector $(W_1^N, \dots, W_D^N) \approx (M_1, \dots, M_D) \sim \text{Mult}(k - B, q_1, \dots, q_D)$, with $q_i = \frac{i e_i(t)}{\sum_j j e_j(t)}$.

If $v \in \mathcal{E}_n^N(0)$, then $Y_{n+1}^N - Y_n^N = (1, 0, -1, 0, \dots, 0)$. □

Proposition 3.3. *The cost function $L(\mathbf{x}(t), \dot{\mathbf{x}}(t))$ defined in (3.1) coincides with the LDP rate function for the empirical average of i.i.d random variables $(Z_i^{\mathbf{x}(t)})_{i \in \mathbb{N}}$ distributed as $Z^{\mathbf{x}(t)}$:*

$$L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = \sup_{\alpha} \{ \langle \alpha, \dot{\mathbf{x}}(t) \rangle - H(\mathbf{x}(t), \alpha) \} = \Lambda_{Z^{\mathbf{x}(t)}}^*(\dot{\mathbf{x}}(t)).$$

Proof. Assuming that the $(Z_i^{\mathbf{x}(t)})_i$ are i.i.d., Cramér's theorem states that the LDP rate function for their empirical average is $I(x) = \Lambda_{Z^{\mathbf{x}(t)}}^*(x) = \sup_{\alpha \in \mathbb{R}^{D+3}} \{ \langle \alpha, x \rangle - \Lambda_{Z^{\mathbf{x}(t)}}(\alpha) \}$,

with $\Lambda_{Z^{\mathbf{x}(t)}}(\alpha) = \log \mathbb{E} \left[e^{\langle \alpha, Z^{\mathbf{x}(t)} \rangle} \right]$.

In this case, with $\alpha = (\alpha_s, \alpha_u, \alpha_0, \dots, \alpha_D)$, we have $\Lambda_{Z^{\mathbf{x}(t)}}(\alpha) = H(\mathbf{x}(t), \alpha)$, being $H(x, \alpha)$ ($H : M \times \mathbb{R}^{D+3} \rightarrow \mathbb{R}$) the log of the moment-generating function of the (conditioned) multinomial vector Z^x (with $x \in M$) evaluated in α , which is presented in Equation (3.2). Defining $L : M \times \mathbb{R}^{D+3} \rightarrow \mathbb{R}$ as in Equation (3.1), results that $L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = \sup_{\alpha} \{ \langle \alpha, \dot{\mathbf{x}}(t) \rangle - H(\mathbf{x}(t), \alpha) \}$ coincides with $\Lambda_{Z^{\mathbf{x}(t)}}^*(\dot{\mathbf{x}}(t))$. □

That is, the global cost of a deviation of $\{X_t^N\}_t$ to a trajectory $\mathbf{x}(t)$ can be interpreted as a consequence of the accumulated cost of microscopic deviations of the average of (conditioned) multinomial random vectors, representing the degrees of the new explored vertices in one step.

3.2 Fluid limit of the process $\{X_t^N\}_t$

In this subsection, we deduce the fluid limit of the process $\{X_t^N\}_t$.

Proposition 3.4 (Fluid limit). *The sequence of processes $\{X^N\}_N$ converges almost-sure, as $N \rightarrow \infty$, to the deterministic function $\hat{x} : [0, 1] \rightarrow M$ given by*

$$\hat{x}(t) = \begin{cases} (s(t), \hat{u}(t), \hat{e}_0(t), \dots, \hat{e}_D(t)), & \text{if } t \leq T^*, \\ (T^*, 0, \dots, 0), & \text{if } t > T^*, \end{cases} \quad \text{where } \hat{e}_i(t) = \begin{cases} e_i(t), & \text{if } t \leq t_i, \\ 0, & \text{if } t > t_i. \end{cases}$$

The times t_i are defined by $t_i = \inf\{t \in [0, 1] : e_i(t) \leq 0\}$ where $x(t) = (s(t), u(t), e_0(t), \dots, e_D(t))$ is (the) solution of the following ordinary differential equation:

$$\begin{cases} \dot{s} = 1, \\ \dot{u} = \frac{-2 \sum_k k e_k}{\sum_k e_k}, \\ \dot{e}_i = \frac{-e_i - \frac{i e_i}{u} \sum_k k e_k}{\sum_k e_k}, \quad i = 0, \dots, D, \\ s(0) = 0, u(0) = \lambda, e_i(0) = p_i. \end{cases} \quad (3.3)$$

\hat{u} is the solution of Equation (3.3) replacing e_i by \hat{e}_i and $T^* = \inf\{t \in [0, 1] : \sum_k \hat{e}_k(t) = 0\} = \max\{t_1, \dots, t_D\}$.

Proof. The cost function $L(x, \beta)$ defined in Theorem 3.1 satisfies $L(x, \beta) \geq 0$ and $L(x, \beta) = 0$ if and only if $\beta = H_\alpha(x, 0)$, where $H_\alpha(x, \alpha)$ are the partial derivatives of $H(x, \alpha)$ w.r.t. $\alpha = (\alpha_s, \alpha_u, \alpha_0, \dots, \alpha_D)$. Then, the trajectories with zero cost are the ones that verify $\dot{x} = H_\alpha(x, 0)$. If in addition we impose the condition $x(0) = (0, \lambda, p_0, \dots, p_D)$, we obtain the autonomous Equation (3.3). Cauchy-Peano existence theorem ensures the existence of at least one solution of such equation. Let $\mathcal{D} = \{x \in M : e_i > 0 \forall i\}$ and $f(x) = H_\alpha(x, 0)$. Then f is a C^1 -function on \mathcal{D} , i.e. it is a locally Lipschitz continuous function on \mathcal{D} . This implies the uniqueness of solutions $e_i(t)$ for equation

$\begin{cases} \dot{x} = f(x), \\ x(0) = x_0 \in \mathcal{D}, \end{cases}$ until the time t_i at which $e_i(t_i) = 0$, and then we take the solution $e_i(t) = 0$ for all $t \geq t_i$. □

3.3 Optimization of the rate function

The following proposition allows to transform the optimization problem of the rate function I over a set of trajectories into a much simpler optimization problem on \mathbb{R}^{D+3} .

Proposition 3.5 (Rate function optimization). *Let A be a subset of $D_M[0, 1]$. Then,*

$$\inf_{x \in A} I(x) = \inf_{\{\alpha_0 \in \mathbb{R}^{D+3} : \hat{x}_{\alpha_0} \in \bar{A}\}} I(\hat{x}_{\alpha_0}),$$

where the closure of A is considered w.r.t. the Skorohod topology,

$$\begin{aligned} \hat{x}_{\alpha_0}(t) &= \begin{cases} (s_{\alpha_0}(t), \hat{u}_{\alpha_0}(t), \hat{e}_{0,\alpha_0}(t), \dots, \hat{e}_{D,\alpha_0}(t)), & \text{if } t \leq T_{\alpha_0} \\ (T_{\alpha_0}, 0, \dots, 0), & \text{if } t > T_{\alpha_0} \end{cases}, \\ \hat{e}_{i,\alpha_0}(t) &= \begin{cases} e_{i,\alpha_0}(t), & \text{if } t \leq t_{i,\alpha_0}, \\ 0, & \text{if } t > t_{i,\alpha_0} \end{cases}, \quad t_{i,\alpha_0} = \inf\{t \in [0, 1] : e_{i,\alpha_0}(t) \leq 0\}, \text{ and} \\ x_{\alpha_0}(t) &= (s_{\alpha_0}(t), u_{\alpha_0}(t), e_{0,\alpha_0}(t), \dots, e_{D,\alpha_0}(t)) \end{aligned}$$

is (the) solution of the ODE:

$$\begin{cases} \dot{x} = H_x(x, \alpha); \\ \dot{\alpha} = -H_\alpha(x, \alpha); \\ x(0) = (0, \lambda, p_0, \dots, p_D); \alpha(0) = \alpha_0. \end{cases} \quad (3.4)$$

H_x and H_α are the vectors of partial derivatives of H w.r.t. x and α , and $T_{\alpha_0} = \inf \{t \in [0, 1] : \sum_k \hat{e}_{k, \alpha_0}(t) = 0\}$.

Remark 3.6. As expected, for $\alpha_0 = (0, 0, \dots, 0)$, $\hat{x}_{\alpha_0}(t)$ coincides with the fluid limit, which is solution of Equation (3.3), and $\alpha(t) = (0, \dots, 0)$ for all t . Then $\inf_{x \in A} I(x) = 0$ if the fluid limit belongs to A .

Proof. Note that if $\mathbf{x} \in \mathcal{H}_L$ is such that $\mathbf{x}(t) = (s(t), u(t), e_0(t), \dots, e_D(t))$ and $\sum_k e_k(t) = 0$ for all $t \geq t_0$, then $I(\mathbf{x}) = \int_0^1 L(\mathbf{x}, \dot{\mathbf{x}})dt = \int_0^{t_0} L(\mathbf{x}, \dot{\mathbf{x}})dt$, so just consider Hamilton's equations for the case $\sum_k e_k > 0$. Hamilton's equations, presented in Equation (3.4), give conditions for a function \mathbf{x} to be a *stationary curve* of the functional I (and are equivalents to Euler-Lagrange equation, see [1], for example). Note that α is an auxiliary function. \square

3.4 LD for the size of the independent set constructed by the algorithm

We can now deduce LD results for the sequence of stopping times $\frac{T_N^*}{N}$ (which coincide with the proportion of vertices in the independent set constructed by the algorithm) using our previous results:

Theorem 3.7. Consider T_N^* defined before as the stopping time of the algorithm presented in Sect. 2, and let $T^* \in [0, 1]$ the time corresponding to the fluid limit defined in Proposition 3.4.

1. If $\varepsilon > 0$ is such that $T^* + \varepsilon < 1$, then

$$\lim_{\frac{1}{N}} \log \mathbb{P} \left(\frac{T_N^*}{N} \geq T^* + \varepsilon \right) = -F^+(T^* + \varepsilon),$$

being $F^+(T^* + \varepsilon) = \inf \{I(\hat{x}_{\alpha_0}) : T_{\alpha_0} \geq T^* + \varepsilon, \alpha_0 \in \mathbb{R}^{D+3}\}$.

2. If $\varepsilon > 0$ is such that $T^* - \varepsilon > 0$, then

$$\lim_{\frac{1}{N}} \log \mathbb{P} \left(\frac{T_N^*}{N} \leq T^* - \varepsilon \right) = -F^-(T^* - \varepsilon),$$

being $F^-(T^* - \varepsilon) = \inf \{I(\hat{x}_{\alpha_0}) : T_{\alpha_0} \leq T^* - \varepsilon, \alpha_0 \in \mathbb{R}^{D+3}\}$.

In both cases \hat{x}_{α_0} and T_{α_0} are as in Proposition 3.5.

Proof. We only prove the first statement because the proof of the second one is analogous. Define the set A_ε , that contains the trajectories $\mathbf{x} \in D_M[0, 1]$ such that $\mathbf{x}(t) = (s(t), u(t), e_0(t), \dots, e_D(t))$, $\mathbf{x}(0) = (0, \lambda, p_0, \dots, p_d)$, coordinates $e_i(t)$, $u(t)$ are decreasing, $s(t)$ is increasing, $0 \leq e_i(t), s(t) \leq 1$ for all t , and such that $\inf \{t : \sum_k e_k(t) = 0\} \geq T^* + \varepsilon$. Then, Proposition 3.5 implies that

$$\begin{aligned} \lim_{\frac{1}{N}} \log \mathbb{P} \left(\frac{T_N^*}{N} \geq T^* + \varepsilon \right) &= \lim_{\frac{1}{N}} \log \mathbb{P} (X_N^N \in A_\varepsilon) \\ &= - \inf_{\{\alpha_0 \in \mathbb{R}^{D+3} : \hat{x}_{\alpha_0} \in \bar{A}_\varepsilon\}} I(\hat{x}_{\alpha_0}) = F^+(T^* + \varepsilon). \quad \square \end{aligned}$$

4 d -regular case

In this section, we analyze the results presented in previous sections for the particular case of a d -regular graph, i.e. $p_d = 1$ and $p_i = 0$ for all $i \neq d$. In this case, the sequence of interest is $\{X_t^N\}_{t \in [0,1]}$ with $X_t^N = \frac{1}{N}Y_{[Nt]}^N$, being $Y_n^N = (S_n^N, U_n^N, E_n^N)$, and

- $S_n^N = |S_n^N|$, the number of vertices that have already been assigned to the independent set at step n ;
- U_n^N , the total number of unpaired half-edges at step n ;
- $E_n^N = |\mathcal{E}_n^N|$, the number of empty vertices at step n .

$X_t^N \in M^N$, being $M^N = \{\frac{1}{N}(\hat{s}, \hat{u}, \hat{e}) : \hat{s}, \hat{e} \in \{0, \dots, N\}; \hat{u} \in \{0, \dots, dN\}; \hat{u} \geq d\hat{e}\}$, which is a subset of the compact set $M := \{(x_1, x_2, x_3) \in [0, 1] \times [0, d] \times [0, 1] : x_2 \geq dx_3\} \subset \mathbb{R}^3$. The Hamiltonian $H : M \times \mathbb{R}^3 \rightarrow \mathbb{R}$ is given by

$$H(x, \alpha) = \begin{cases} \alpha_1 - 2d\alpha_2 - \alpha_3 + d \log \left[1 + (e^{-\alpha_3} - 1) \frac{dx_3}{x_2} \right], & \text{if } x_3 > 0, \\ 0, & \text{if } x_3 = 0, \end{cases} \tag{4.1}$$

where $x = (x_1, x_2, x_3)$, $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, and the cost function L can be obtained explicitly as $L : M \times \mathbb{R}^3 \rightarrow \mathbb{R}$ such that

$$L(x, \beta) = \begin{cases} (\beta_3 + 1)\alpha_3^* - d \log \left[1 + (e^{-\alpha_3^*} - 1) \frac{dx_3}{x_2} \right], & \text{with } \alpha_3^* = \log \left[\frac{dx_3}{dx_3 - x_2} \left(\frac{d}{\beta_3 + 1} + 1 \right) \right], \\ & \text{if } \beta_1 = 1, \beta_2 = -2d, \beta_3 \geq -(d + 1), \\ 0, & \text{if } x_3 = \beta_3 = 0, \\ +\infty, & \text{in other cases.} \end{cases} \tag{4.2}$$

4.1 Fluid limit

The trajectories with zero cost are $\mathbf{x}(t) = (s(t), u(t), e(t))$ such that $\dot{s} = 1$, $\dot{u} = -2d$, and $\dot{e} = -1 - \frac{d^2 e}{u}$. For the initial condition $\mathbf{x}(0) = (0, d, 1)$ with $d \geq 3$, the unique solution is $\mathbf{x}(t) = (s(t), u(t), e(t))$ with $s(t) = t$, $u(t) = d(-2t + 1)$, and

$$e(t) = \begin{cases} \frac{1}{d-2} \left[2t - 1 + (d-1)(1-2t)^{\frac{d}{2}} \right], & \text{if } t \leq T^*, \\ 0, & \text{if } t > T^*, \end{cases}$$

where $T^* = \inf \{t : e(t) = 0\}$, i.e. the *jamming constant* is $T^* = \frac{1}{2} \left[1 - \left(\frac{1}{d-1} \right)^{\frac{2}{d-2}} \right]$. This coincides with the known result given in [15].

If $d = 2$, then the fluid limit is $\mathbf{x}(t) = (s(t), u(t), e(t))$ with the same functions $s(t)$ and $u(t)$, and $e(t)$ is given by

$$e(t) = \begin{cases} (1 - 2t) \left[\frac{1}{2} \log(1 - 2t) + 1 \right], & \text{if } t \leq T^*, \\ 0, & \text{if } t > T^*, \end{cases}$$

where $T^* = \frac{1-e^{-2}}{2}$. This coincides with the known result from the earlier work of [11].

4.2 LDP for the number of unexplored vertices

Since the trajectories with positive probability for X_t^N , when $N \rightarrow \infty$, are those $\mathbf{x}(t) = (s(t), u(t), e(t))$ such that $s(t) = t$ and $u(t) = d(-2t + 1)$, we can directly deduce an LDP for a rescaling of the process that counts the number of unexplored vertices in each step of the algorithm. With abuse of notation, let $\{E_t^N\}_t$ be a rescaling of the process that counts the number of unexplored vertices in each step given by $E_t^N := \frac{E_{[Nt]}^N}{N}$.

LD for the greedy exploration process on configuration models

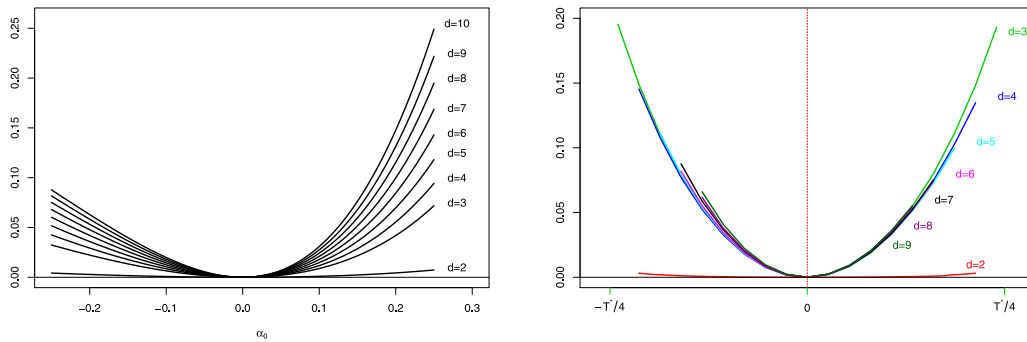


Figure 1: *Left:* Evolution of $\hat{F}(\alpha_0)$ as function of α_0 for $d = 2, \dots, 10$. *Right:* Evolution of $\hat{F}(\alpha_0(T^* \pm \varepsilon))$ as function of ε for $d = 2, \dots, 9$.

Corollary 4.1. *The sequence of processes $\{E^N\}_N$ given by $E_t^N := \frac{E_{[Nt]}^N}{N}$ verifies an LDP in $D_{[0,1]}[0, 1]$ with good rate function $\hat{I} : D_{[0,1]}[0, 1] \rightarrow [0, +\infty]$ such that $\hat{I}(x) = \int_0^1 \hat{L}(t, x(t), \dot{x}(t)) dt$, where $\hat{L}(t, x, y) = L((t, u(t), x), (1, -2d, y))$, with $u(t) = d(-2t + 1)$.*

Proof. It is deduced directly from Theorem 3.1. □

Remark 4.2. The cost function $\hat{L}(t, x, y)$ verifies $\hat{L}(t, x, y) = \Lambda_{W_{t,x}}^*(y)$, being $\Lambda_{W_{t,x}}^*(y)$ the LD rate function for the average of random variables $W_{t,x} = B_{t,x} - d - 1$, where $B_{t,x}$ has a binomial distribution $\text{Bin}(d, p = 1 - \frac{dx}{u(t)})$. This corresponds to the approximation to the distribution of new explored vertices in each step of the algorithm. Note that $\Lambda_{W_{t,x}}^*(y) = \Lambda_{B_{t,x}}^*(y + d + 1)$.

In this case, we can explicitly deduce the LD rate for the size of the independent set constructed by the algorithm since the optimization problem of the rate function \hat{I} over a set of trajectories of $D_{[0,1]}[0, 1]$ becomes an optimization problem in \mathbb{R} .

Corollary 4.3 (Optimization of the rate function \hat{I}). *Let A be a subset of $D_{[0,1]}[0, 1]$. Then, $\inf_{x \in A} \hat{I}(x) = \inf_{\{\alpha_0 \in \mathbb{R} : \hat{x}_{\alpha_0} \in A\}} \hat{F}(\alpha_0)$, being $\hat{F}(\alpha_0) = \int_0^{T_{\alpha_0}} \hat{L}(t, x_{\alpha_0}(t), \dot{x}_{\alpha_0}(t)) dt$, x_{α_0} is the solution of the ODE:*

$$\begin{cases} \dot{x} = -1 + \frac{dx}{e^{y(2t-1+x)} - x}; \\ \dot{y} = \frac{d(1-e^y)}{e^{y(2t-1+x)} - x}; \\ x(0) = 1; y(0) = \alpha_0, \end{cases} \tag{4.3}$$

$$T_{\alpha_0} = \inf \{t \in [0, 1] : x_{\alpha_0} \leq 0\}, \text{ and } \hat{x}_{\alpha_0}(t) = \begin{cases} x_{\alpha_0}(t), & \text{if } 0 \leq t \leq T_{\alpha_0}, \\ 0, & \text{if } t > T_{\alpha_0}. \end{cases}$$

Proof. It is a corollary of Proposition 3.5. □

In Fig. 1 (left), the evolution of \hat{F} as a function of α_0 is presented for $d = 2, \dots, 10$. It is observed that when $d = 2$ (where the problem basically boils down to finding independent sets on circles), the possibility of deviating from the fluid limit is much higher than for larger d .

Corollary 4.4. *The sequence of stopping times T_N^* defined over a d -regular graph $G(N, d)$ verifies:*

1. If $\varepsilon > 0$ is such that $T^* + \varepsilon < 1$, then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left(\frac{T_N^*}{N} \geq T^* + \varepsilon \right) = -\hat{F}(\alpha_0(T^* + \varepsilon)),$$

where $\alpha_0(T^* + \varepsilon)$ is the unique real number $\alpha_0 > 0$ such that $T_{\alpha_0} = T^* + \varepsilon$.

2. If $\varepsilon > 0$ is such that $T^* - \varepsilon > 0$, then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left(\frac{T_N^*}{N} \leq T^* - \varepsilon \right) = -\hat{F}(\alpha_0(T^* - \varepsilon)),$$

where $\alpha_0(T^* - \varepsilon)$ is the unique real number $\alpha_0 < 0$ such that $T_{\alpha_0} = T^* - \varepsilon$.

In both cases $\hat{F}(\alpha_0)$ and T_{α_0} are as in Proposition 4.3.

Figure 1 (right) presents the evolution of $\hat{F}(\alpha_0(T^* \pm \varepsilon))$ as a function of $\varepsilon \in \left[0, \frac{T^*}{4}\right]$ for $d = 3, \dots, 9$, compared with $\varepsilon \in \left[0, \frac{T^*}{6}\right]$ for $d = 2$. Note that in each case the time T^* depends on d . Again, the abrupt change in the dynamics is observed for $d = 2$ and $d > 2$.

Proof. In this case, if x_{α_0} is the solution of Equation (4.3) with $y(0) = \alpha_0$, then the following monotony property with respect to the initial condition α_0 is verified: if $\alpha_0 < \alpha_1 \Rightarrow x_{\alpha_0}(t) < x_{\alpha_1}(t)$ for all t , then $T_{\alpha_0} < T_{\alpha_1}$. In addition, it can be seen that for all $T \in (T^*, 1)$, there exists a unique value $\alpha_0 = \alpha_0(T) > 0$ such that $x_{\alpha_0}(T) = 1$ (i.e. $T = T_{\alpha_0}$). Then, there is only one $\alpha_0^* > 0$ such that $x_{\alpha_0^*}(T^* + \varepsilon) = 1$. If A_ε is the set defined in the proof of Theorem 3.7, then $\hat{x}_{\alpha_0} \in A_\varepsilon \Leftrightarrow \alpha_0 \geq \alpha_0^*$, which implies that

$$\inf_{\{\alpha_0: \hat{x}_{\alpha_0} \in A_\varepsilon\}} \hat{F}(\alpha_0) = \inf_{\{\alpha_0 \geq \alpha_0^*\}} \hat{F}(\alpha_0) = \hat{F}(\alpha_0^*). \quad \square$$

5 Proof of Theorem 3.1

The proof of Theorem 3.1 goes similar to the one included with full details in [3] for the process that counts the number of unexplored vertices of an exploration algorithm over Erdős-Renyi graphs. Basically, it consists in verifying that the sequence of processes $\{X^N\}_N$ defined in Sect. 2 verifies the assumptions from [9], which are organized bellow as steps 1,2,3, 4. Once these hypotheses have been verified, Theorem 3.1 results from using Theorems 5.15, 8.14, and 8.23 from [9]. In a nutshell, as a consequence of the first two steps, the process $\{X^N\}_N$ verifies the exponential tightness condition. Step 3 assures an LDP via the comparison principle, and finally, Step 4 provides a useful variational version of the rate.

Let T^N be the linear generator of $\left\{\frac{Y_n^N}{N}\right\}_n$, being $\{Y_n^N\}_n$ the discrete-time Markov process defined in Sect. 2. Let $\mathbf{H}_N : \text{Dom}(\mathbf{H}_N) \subset B(M) \rightarrow B(M)$ be the non-linear generator given by $\mathbf{H}_N(f)(x) = \log(e^{-Nf(x)} T^N(e^{Nf})(x))$, where $x \in M^N \subset M$, and $B(M)$ is the space of bounded and Borel measurable functions $f : M \rightarrow \mathbb{R}$.

Step 1: Verify the convergence of the sequence of operators \mathbf{H}_N

The first step consists of proving the convergence of non-linear generators \mathbf{H}_N . In this case, if $x \in M^N \subset M$, then

$$\begin{aligned} T^N(f)(x) &= \mathbb{E} \left[f \left(\frac{Y_{n+1}^N}{N} \right) \middle| \frac{Y_n^N}{N} = x \right] = f \left(x + \frac{1}{N} (1, 0, -1, 0, \dots, 0) \right) \frac{e_0}{\sum_j e_j} \\ &\quad + \sum_{k=1}^D \sum_{h=0}^k \sum_{b=0}^h \sum_{\tilde{w}_j: \sum \tilde{w}_j \leq h-b} f \left(x + \frac{1}{N} (1, -k - h, 0, -\tilde{w}_1, \dots, -1 - \tilde{w}_k, \dots, -\tilde{w}_D) \right) \\ &\quad \times p_N \left(x, k, h, b, (\tilde{w}_j)_j \right). \end{aligned}$$

The probability $p_N(x, k, h, b, (\tilde{w}_j)_j)$ is given by

$$\begin{aligned} & p_N(x, k, h, b, (\tilde{w}_j)_j) \\ &= \mathbb{P}\left(\left(\tilde{W}_j^N\right)_j = (\tilde{w}_j)_j \mid \frac{Y_n^N}{N} = x; v \in \mathcal{E}_n^N(k); \tilde{H}^N = h; B^N = b; (W_j^N)_j = (\tilde{w}_j)_j\right) \\ &\quad \times \mathbb{P}\left(\left(W_j^N\right)_j = (\tilde{w}_j)_j \mid \frac{Y_n^N}{N} = x; v \in \mathcal{E}_n^N(k); \tilde{H}^N = h; B^N = b\right) \\ &\quad \times \mathbb{P}\left(B^N = b \mid \frac{Y_n^N}{N} = x; v \in \mathcal{E}_n^N(k); \tilde{H}^N = h\right) \\ &\quad \times \mathbb{P}\left(\tilde{H}^N = h \mid \frac{Y_n^N}{N} = x; v \in \mathcal{E}_n^N(k)\right) \times \mathbb{P}\left(v \in \mathcal{E}_n^N(k) \mid \frac{Y_n^N}{N} = x\right), \end{aligned}$$

where $\tilde{H}^N = H^N$ if $k - H^N$ is even and it is $H^N + 1$ otherwise, $H^N \sim \mathbf{Hyper}(Nu, Nu - k, k)$, $B^N \sim \mathbf{Hyper}(Nu - k, Nu - \sum_j jNe_j, \tilde{H}^N)$, and

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(\left(\tilde{W}_j^N\right)_j = (\tilde{w}_j)_j \mid \frac{X_n^N}{N} = x; v \in \mathcal{E}_n^N(k); \tilde{H}^N = h; B^N = b; (W_j^N)_j = (\tilde{w}_j)_j\right) = 1$$

by Lemma 2.1. Then,

$$\begin{aligned} & e^{-Nf(x)} T^N(e^{Nf})(x) \\ &= e^{N(f(x + \frac{1}{N}(1,0,-1,0,\dots,0)) - f(x))} \frac{e_0}{\sum_j e_j} \\ &\quad + \sum_{k=1}^D \sum_{h=0}^k \sum_{b=0}^h \sum_{\tilde{w}_j: \sum \tilde{w}_j \leq h-b} e^{N(f(x + \frac{1}{N}(1,-k-h,0,-\tilde{w}_1,\dots,-1-\tilde{w}_k,\dots,-\tilde{w}_d)) - f(x))} \\ &\quad \times p_N(x, k, h, b, (\tilde{w}_j)_j). \end{aligned}$$

Using Taylor’s theorem and Stirling’s formula we prove² that \mathbf{H}_N converges to \mathbf{H} in the following sense: $\lim_{N \rightarrow \infty} \sup_{x \in M^N} |\mathbf{H}_N(f)(x) - \mathbf{H}(f)(x)| = 0$ for all $f \in C^1(M)$, being $\mathbf{H} : C^1(M) \rightarrow \mathbb{R}$ such that $\mathbf{H}(f)(x) = H(x, \nabla f(x))$, and H is defined in Equation (3.2).

Step 2: Verify the exponential compact containment condition

The second step consists of verifying the *exponential compact containment condition* (see Condition 2.8 from [9]). In our case, it is trivially verified since the state space M is a compact subset of \mathbb{R}^{D+3} .

Step 3: Prove that \mathbf{H} generates a semigroup $\mathbf{V} = \{V_t\}_t$

The third step consists of proving that \mathbf{H} generates a semigroup $\mathbf{V} = \{V_t\}_t$. This is the most technical step. \mathbf{H} generates a semigroup if there is $\mathbf{V} = \{V_t\}_t$ such that for all $f \in \text{Dom}(\mathbf{V})$,

$$\frac{d}{dt} V_t(f) = \mathbf{H}(V_t(f)); \quad V_0(f) = f.$$

The main theorem of [5] implies that $\mu_N(t) = (Id - \frac{t}{N} \mathbf{H})^{-N}$ converges to the solution of the previous equation if \mathbf{H} is m-dissipative. Then we need to prove that for all $h \in C(M)$

²First for the case $f \in C^2(M)$ and then it is extended to $f \in C^1(E)$ by taking a sequence $\{f_m\}_m \subset C^2(M)$ such that $\limsup_m \sup_{x \in M} |f_m(x) - f(x)| = 0$ and the triangular inequality.

and $\beta > 0$, there exists $f \in C^1(M)$ such that

$$f - \beta \mathbf{H}(f) - h = 0. \tag{5.1}$$

However the verification of this property can be a formidable obstacle. One way out is to work with viscosity solutions and prove that the *comparison principle* for Equation (5.1) is verified. If the comparison principle is verified, then the operator \mathbf{H} can be extended to $\hat{\mathbf{H}}$ such that $\hat{\mathbf{H}}$ is m-dissipative and generates a semigroup \mathbf{V} (see Theorem 8.27 of [9]).

Then, since in this case the operator \mathbf{H} can be written as $\mathbf{H}(f)(x) = H(x, \nabla f(x))$, this step consists of proving that the Hamilton-Jacobi equation:

$$f(x) - \beta H(x, \nabla f(x)) - h(x) = 0, \tag{5.2}$$

has a unique solution f for all $h \in C(M)$ and $\beta > 0$ in a viscosity sense. In our case, we use results from [12] to prove that this Hamilton-Jacobi equation verifies the *comparison principle*, which ensures the uniqueness of the viscosity solution, explicitly constructed in Chap. 9 of [9]. As a consequence of Proposition 4.2 in [12], it is enough to prove that the following inequality holds:

$$\liminf_{\alpha \rightarrow \infty} H(x^\alpha, \alpha \psi_x(x^\alpha, y^\alpha)) - H(y^\alpha, \alpha \psi_x(x^\alpha, y^\alpha)) \leq 0, \tag{5.3}$$

where $\psi_x(x, y)$ is the vector of partial derivatives w.r.t. $x = (s, u, e_0, \dots, e_D)$ of the good distance function $\psi(x, y) = \frac{1}{2} \|x - y\|^2$. $x^\alpha = (s^{x^\alpha}, u^{x^\alpha}, e_0^{x^\alpha}, \dots, e_D^{x^\alpha})$ and $y^\alpha = (s^{y^\alpha}, u^{y^\alpha}, e_0^{y^\alpha}, \dots, e_D^{y^\alpha})$ are the sequences constructed in Chap. 9 of [9] (with $\alpha \rightarrow +\infty$) and verify $(x^\alpha, y^\alpha) \rightarrow (z, z)$ where $\mu(z) - v(z) = \sup_{x \in M} \{\mu(x) - v(x)\}$ for a given (viscosity) subsolution μ and supersolution v of Equation (5.2).

Step 4: Construct a variational representation for the rate function I

Finally, the limiting semigroup $\mathbf{V} = \{V_t\}_t$ usually admits a variational form $\{\mathcal{V}_t\}_t$, known as the *Nisio semigroup* in control theory. Then, the fourth step consists of providing the more treatable form of the rate function I presented in Theorem 3.1. In our case, as $\mathbf{H}(f)(x) = H(x, \nabla f(x))$ for each $x \in M$ and $H \leftrightarrow L$,³ we have that \mathbf{H} can be written as $\mathbf{H}(f)(x) = \sup_{u \in U} \{A(f)(x, u) - L(x, u)\}$, where $U = \mathbb{R}^{D+3}$ and A is the linear operator given by $A(f)(x, u) = \langle \nabla f(x), u \rangle$ for $f \in C^1(M)$ and $(x, u) \in M \times U$. Then, we consider the Nisio semigroup corresponding to the control problem determined by A and the cost function $-L$:

$$\mathcal{V}_t(f)(x_0) = \sup_{\{(x, \lambda) \in \mathcal{Y}: x(0) = x_0\}} \left\{ f(\mathbf{x}(t)) - \iint_{U \times [0, t]} L(\mathbf{x}(s), u) \lambda(du \times ds) \right\} \tag{5.4}$$

for each $x_0 \in M$. $\mathcal{Y} \subset D_M[0, 1] \times \mathcal{M}_m(U)$ is the *control set* of the linear operator A (see Definition 8.1 from [9]), and $\mathcal{M}_m(U)$ is the space of Borel measures λ on $U \times [0, 1]$ satisfying $\lambda(U \times [0, t]) = t$ for all $t \in [0, 1]$. Measure λ is known as a *relaxed control*. As L is convex w.r.t. β , it follows that a deterministic control $\lambda(du \times ds) = \delta_{u(s)}(du)ds$ is always the control with smallest cost by Jensen’s inequality, and the supremum in Equation (5.4) is attained on $\mathcal{Y}_{AC} := \{(x, \lambda) \in \mathcal{Y} : x \in \mathcal{AC}, x(0) = x_0\}$, being \mathcal{AC} the space of absolutely continuous functions $x : [0, 1] \rightarrow M$.

Then, as consequence of Theorems 8.14 and 8.23 from [9], it is enough to prove that Conditions 8.9, 8.10 and 8.11 from [9] are verified to prove that $V_t = \mathcal{V}_t$ and I can be written as in Theorem 3.1.

³We use $H \leftrightarrow L$ to denote that $L(x, \beta) = \sup_\alpha \{\langle \alpha, \beta \rangle - H(x, \alpha)\}$ and $H(x, \alpha) = \sup_\beta \{\langle \alpha, \beta \rangle - L(x, \beta)\}$.

To prove these conditions, we consider elements (x, λ) from \mathcal{Y}_{AC} and use that $L(x, q(x)) = 0$ for all $x \in M$ if $q(x) = H_\alpha(x, 0)$, and $H(x, \nabla f(x)) = \langle \nabla f(x), q_f(x) \rangle - L(x, q_f(x))$ if $q_f(x) = H_\alpha(x, \nabla f(x))$.

References

- [1] Arnold, B. I.: Métodos Matemáticos da Mecânica Clásica. *Editora Mir Moscovo* (1987).
- [2] Bermolen, P., Jonckheere, M. and Moyal, Pascal: The jamming constant of uniform random graphs. *Stochastic Processes and Their Applications* **7** (2017), 2138–2178. MR3652409
- [3] Bermolen, P., Goicoechea, V., Jonckheere, M. and Mordecki, E.: Large deviation principle for the greedy exploration algorithm over Erdős-Rényi graphs. *ALEA, Lat. Am. J. Probab. Math. Stat.* **19** (2022), 439–456. MR4359798
- [4] Brightwell, Graham and Janson, Svante and Luczak, Malwina: The greedy independent set in a random graph with given degrees. *Random Structures & Algorithms* **51** (2017), 565–586. MR3718590
- [5] Crandall, M. G. and Liggett, T. M.: Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.* **93** (1971), 265–298. MR287357
- [6] de Acosta, A.: Exponential tightness and projective systems in large deviation theory. *Festschrift for Lucien Le Cam* (1997), 143–156. MR1462943
- [7] Dupuis, Paul and Ellis, Richard S.: A weak convergence approach to the theory of large deviations. *Wiley Series in Probability and Statistics: Probability and Statistics* (1997). MR1431744
- [8] Evans, L. C. and Ishii, H.: A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2** (1985), 1–20. MR781589
- [9] Feng, Jin and Kurtz, Thomas G.: Large deviations for stochastic processes. *Mathematical Surveys and Monographs* **131** (2006). MR2260560
- [10] Fleming, Wendell H.: A stochastic control approach to some large deviations problems. *Lecture Notes in Math.* **1119** (1985), 52–66. MR790689
- [11] Flory, Paul J.: Intramolecular reaction between neighboring substituents of vinyl polymers. *Journal of the American Chemical Society* **61** (1939), 1518–1521.
- [12] Kraaij, Richard: Large deviations for finite state Markov jump processes with mean-field interaction via the comparison principle for an associated Hamilton-Jacobi equation. *J. Stat. Phys.* **164** (2016), 321–245. MR3513255
- [13] Puhalskii, A.: The method of stochastic exponentials for large deviations. *Stochastic Process. Appl.* **54** (1994), 45–70. MR1302694
- [14] Van Der Hofstad, R.: Random Graphs and Complex Networks. *Cambridge Series in Statistical and Probabilistic Mathematics* (2017).
- [15] Wormald, Nicholas C.: Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** (1995), 1217–1235. MR1384372

Acknowledgments. The authors would like to thank the referee for their valuable comments that significantly helped us improve our manuscript.