# Large Deviation Principle for the Greedy Exploration Algorithm over Erdös-Rényi Graphs

Paola Bermolen, Valeria Goicoechea, Matthieu Jonckheere, Ernesto Mordecki

# Large Deviation Principle for the Greedy Exploration Algorithm over Erdös-Rényi Graphs

**P. Bermolen, V. Goicoechea, M.Jonckheere and E.Mordecki**

Instituto de Matemática y Estadística Prof. Rafael Laguardia,
Facultad de Inegeniería,
Universidad de la República,
Uruguay.
*E-mail address*: paola@fing.edu.uy

Instituto de Matemática y Estadística Prof. Rafael Laguardia,
Facultad de Inegeniería,
Universidad de la República,
Uruguay.
*E-mail address*: vgoicoechea@fing.edu.uy

Instituto de Cálculo. Conicet,
Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires,
Argentina.
*E-mail address*: mjonckhe@dm.uba.ar

Centro de Matemática,
Facultad de Ciencias,
Universidad de la República,
Uruguay.
*E-mail address*: emordecki@cmat.edu.uy

**Abstract.** We prove a large deviation principle for a greedy exploration process on an Erdös-Rényi (ER) graph when the number of nodes goes to infinity. To prove our main result, we use the general strategy to study large deviations of processes proposed by Feng and Kurtz (2006), based on the convergence of non-linear semigroups. The rate function can be expressed in a closed-form formula, and associated optimization problems can be solved explicitly, providing the large deviation trajectory. Also, we derive an LDP for the size of the maximum independent set discovered by such an algorithm and analyze the probability that it exceeds known bounds for the maximal independent set. We also analyze the link between these results and the landscape complexity of the independent set and the exploration dynamic.

## 1. Introduction

Consider a finite, possibly random, graph $G$ for which $V$ is the set of $N$ nodes or vertices. A typical sequential exploration algorithm, usually referred to as *"greedy*

*Key words and phrases.* Large Deviation Principle, Greedy Exploration Algorithms, Erdös-Rényi Graphs, Hamilton-Jacobi equations, Comparison Principle.

Research supported by Stic Amsud Gene.

*algorithm"* [1] works as follows. Initially, all the vertices are declared as *unexplored*. At each step, it selects a vertex and changes its state into *active*. After this, it takes all of its unexplored neighbors and changes their states into *blocked*. The *active* and *blocked* vertices are considered as *explored* and removed from the set of *unexplored* vertices. The algorithm keeps repeating this procedure until step $T_N^*$, at which all vertices are either active or blocked (or equivalently, the set of the *unexplored* vertex is empty). Observe that at any step $k$, the active vertices form an independent set (i.e. there are no edges between the nodes of this set) and that $T_N^*$ is the size of the independent set constructed by the algorithm. Let $Z_k^N$ be the number of explored nodes at time $k$, then $Z_{T_N^*}^N = N$.

Our motivation to study such an exploration process on random graphs is twofold. On the one hand, exploration processes have received a great amount of attention in spatial structures. It has been considered on discrete structures like $\mathbb{Z}^d$ (see Ritchie (2006); Ferrari et al. (2002)) and point processes (see Penrose (2001); Baccelli and Tien Viet (2012)). In physics and biological sciences, where it is usually referred to as random sequential absorption, it models phenomena of deposition of colloidal particles or proteins on surfaces (see Evans (1993)). In communication sciences and wireless networks in particular, it allows to represent the number of connections for CSMA-like algorithms in a given time-slot, for a given spatial configuration of terminals (see Kleinrock and Takagi (1985) for a classical reference on the protocol definition).

On the other hand, these dynamics are the simplest procedure to construct (maximal) independent sets and have been extensively studied for specific graphs. Explicit results for the size of these sets have been obtained for regular graphs in Wormald (1995), exploiting their particular structure; see also Gamarnik and Sudan (2017) for graphs with large girths, and Bermolen et al. (2017b) for more general configuration models. In this context, the greedy algorithm is the simplest instance of a local algorithm, i.e., an algorithm using only local information available at each vertex and using some randomness. Recently, it was proven in Gamarnik and Sudan (2017) that contrary to previously stated conjectures (for instance, in Hatami et al. (2014)), local algorithms can not discover asymptotically maximum independent sets (independent set of the maximum size) and stay sub-optimal, up to multiplicative constant, for regular graphs with large girth. Hence, it is natural to look at related questions for Erdös-Rényi (ER) graphs: we focus on giving estimates of reaching a given size of maximum independent sets by studying the large deviations of the exploration process.

Thanks to the great amount of independence and symmetry of the edges' collection in a sparse ER graph $G(N, c/N)$, the greedy exploration algorithm is characterized by $\left\{Z_k^N\right\}_k$, a simple one-dimensional Markov process. Consequently, a functional law of large numbers described by a differential equation can be employed to get the macroscopic size of the constructed independent set when the number of nodes goes to infinity (see Bermolen et al. (2017a) and references in McDiarmid (1990)). Diffusion approximations for the process and central limit theorem derived from it for the size $T_N^*$ of the associated independent set are also known, see Bermolen et al. (2017a). Moreover, in Pittel (1982), exponential bounds are proved for the probability that the stopping times $t_f(G(N, p/N))$ of the $f$-driven

[1] called greedy although there is no policy to choose the optimal vertex in each step, see for instance the definition of an unweighted greedy algorithm in Jungnickel (2005).

algorithms (in particular, $T_N^*$) belong to certain intervals. However, to the best of our knowledge, there is no characterization of a large deviation principle (LDP) for both the discrete-time Markov process $\left\{Z_k^N\right\}_k$ and the random variable $T_N^*$, which can give various types of useful information both on the greedy exploration and on the independent set landscape. For example, it allows determining the most probable trajectory for which the independent set's size is bigger/samaller than selected bounds. The present paper's topic is a refined analysis of this simple algorithm by studying the large deviations (LD) for the sequence of processes $\left\{Z_k^N\right\}_k$. As a corollary, we obtain an LDP for the size of the independent set constructed by the algorithm.

Although $\left\{Z_k^N\right\}_k$ is a simple Markov process, as far as we know, computing its LDP does not directly follow from classical results. Indeed, the well-known work of Freidlin and Wentzell (1984) is not directly applicable to our process since both the drift and the jump measure involved in the underlying stochastic differential equation depend on the scaling parameter. An LD upper bound for a general family of processes, including processes whose (discontinuous) drift and jump measure depends on the scaling, is presented in Dupuis et al. (1991). However, the authors do not provide sufficient conditions to ensure that the general upper bound obtained for simpler processes is still valid for this case.

In this article, we use techniques from the theory of viscosity solutions to Hamilton-Jacobi equations and prove that its LD upper bound not only works for a continuous-time version of $\left\{Z_k^N\right\}_k$, but is also effectively the LD rate function. To prove this LDP, we use the general strategy to study of large deviations of processes proposed by Feng and Kurtz (2006), which is based on the convergence of non-linear semigroups.

In general, there are at least two approaches in the literature to prove an LDP. The traditional approach to LDP is via the so-called *change of measure method*. Indeed, beginning with the work of Cramér (1938) and including the fundamental work on large deviations for stochastic processes by Freidlin and Wentzell (1984) and Donsker and Varadhan (1975), much of the analysis has been based on a change of measure techniques. In this approach, a tilted or reference measure is identified under which the events of interest have a high probability. The probability of the event under the original measure is bounded in terms of the Radon-Nikodym density that relates both measures. In our case, finding a direct change of measure turns out to be a highly non-trivial task due to the transitions rate dependence on the state and the intricate overall dependence on the scaling parameter.

Another approach is analogous to the Prohorov compactness approach to weak convergence of probability measures (by studying these measures' tightness). It is sometimes referred to as the *exponential tightness method*. This has been established by Puhalskii (1994), O'Brien and Vervaat (1995), de Acosta (1997), Dupuis and Ellis (1997), Fleming (1985), Evans and Ishii (1985), and others.

The remarkable work of Feng and Kurtz (2006) consists of combining the tools of probability, analysis, and control theory used in the works of de Acosta (1997), Dupuis and Ellis (1997), Evans and Ishii (1985), Fleming (1977/78), Fleming (1985), Fleming (1999), Puhalskii (1994), 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 $H_N$ and derive 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 $\mathbf{V} = \{V_t\}_t$. This issue is nontrivial and follows, for example, by showing that the Hamilton-Jacobi equation $f(x) - \beta H\left(x, \nabla f(x)\right) - h(x) = 0$ has a unique solution $f$ for all $h \in C(E)$ and $\beta > 0$ in a viscosity sense when $\mathbf{H}(f)(x) = H\left(x, \nabla f(x)\right)$. The rate function is constructed in terms of that limit $\mathbf{V}$. This limiting semigroup usually admits a variational form known as the *Nisio semigroup* in control theory. Then, the fourth step consists of constructing a variational representation for the rate function. In a nutshell, as a consequence of the first two steps, the process verifies the exponential tightness condition; the third step assures the existence of an LDP, and the fourth step provides a useful variational version of the rate.

In our case, after working on the four steps that we mentioned before, we deduce not only a variational form of the rate function but also prove that it can be expressed as an action integral of a cost function $L$. Moreover, by solving the associated Hamilton's equations, the optimization of the rate over a set of trajectories can be transformed into a real parametric function optimization.

Additionally, the cost function $L$ has a simple interpretation in terms of local deviations for the average of Poisson random variables. As such, this is a first step to understand how such local algorithms behave on complicated landscapes.

This result also allows us to derive quantitative results about the independent set's size constructed by this algorithm. For instance, we can compute the probability that this size is larger than the asymptotic Erdös bound for the maximum independent set when $c \geq 3$ and for the maximum independent set's exact value when $c < e$. In particular, it sheds light on the relation between the complexity of the landscape and the exploration algorithm. It is known (and coined as the $e$-phenomena in Spitzer (1975); Jonckheere and Saenz (2019)) that for $G(N, c/N)$ with $c < e$, an improved local algorithm (the degree-greedy algorithm, which is an improvement of the modification of the greedy algorithm presented in the earlier paper of Karp and Sipser (1981)) is asymptotically optimal. The computation of LD estimates for the greedy exploration (using the asymptotic Erdös bound) allows us to give evidence of a phase transition for the independent set landscape around $e$ (we lose some precision here because of using a bound instead of the true asymptotic value of the independent set), but it hints at an interesting connection between complexity phase transitions and explicit large deviations results.

The rest of the paper is organized as follows. In Section 2, we define our model and present the main result of this article: a path-state LDP for the greedy exploration process. As a corollary, we obtain an LDP for the size of the independent set discovered by the algorithm and analyze its implications. In Section 3, we briefly describe Feng and Kurtz's theory in our context and prove our main theorem.

## 2. Main Results

In this section, we define our process and state the main results of the paper. The key steps of the proof of Theorem 2.1 are presented in the next section.

2.1. *Greedy exploration algorithm.* Let $G\left(N, \frac{c}{N}\right)$ be a sparse Erdös-Rényi graph for which $V$ is the set of $N$ vertices. At any step $k = 0, 1, 2, \ldots$, we consider that each vertex is either *active*, *blocked*, or *unexplored*. Accordingly, the set of vertices

will be split into three components: the set of active vertices $\mathcal{A}_k$, the set of blocked vertices $\mathcal{B}_k$, and the set of unexplored vertices $\mathcal{U}_k$.

The greedy exploration algorithm in discrete time on a graph $G$ can be described as follows. Initially, it sets $\mathcal{U}_0 = V$, $\mathcal{A}_0 = \emptyset$ and $\mathcal{B}_0 = \emptyset$. To explore the graph, at the $(k+1)$-th step it selects uniformly a vertex $i_{k+1} \in \mathcal{U}_k$ and changes its state into active. After this, it takes all of its unexplored neighbors, i.e. the set $\mathcal{N}_{i_{k+1}} = \{w \in \mathcal{U}_k | i_{k+1} \text{ shares and edge with } w\}$, and changes their states into blocked. This means that the resulting set of vertices will be given by $\mathcal{U}_{k+1} = \mathcal{U}_k \backslash \{i_{k+1} \cup \mathcal{N}_{i_{k+1}}\}$, $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{i_{k+1}\}$ and $\mathcal{B}_{k+1} = \mathcal{B}_k \cup \mathcal{N}_{i_{k+1}}$. The algorithm iterates this procedure until the step $T_N^*$ at which all vertices are either active or blocked (or equivalently $\mathcal{U}_{T_N^*} = \emptyset$). Observe that at any step $k$, the active vertices form an independent set and that $\mathcal{A}_{T_N^*}$ is a maximal independent set (because each of the vertices in $V \setminus A_{T_N^*}$ is a neighbour of at least one vertice of $A_{T_N^*}$).

Let $Z_k^N = \left|\mathcal{A}_k^N \cup \mathcal{B}_k^N\right|$ be the number of explored vertices at step $k$. By construction, $Z_{k+1}^N = Z_k^N + 1 + \zeta_{k+1}^N$, where $\zeta_{k+1}^N$ is the number of unexplored neighbors of the selected active vertex at step $k+1$. The distribution of $\zeta_{k+1}^N$ depends only on the number of already explored vertices $Z_k^N$, that is the distribution is Binomial with updated parameter $N - Z_k^N - 1$ and the same edge probability $c/N$. The process $\left\{Z_k^N\right\}_k$ is then a discrete time Markov chain with state space $\{0, 1, 2, ..., N\}$, increasing, time-homogeneous and with an absorbing state $N$. We are interested in $T_N^* \in \{0, 1, 2, ..., N\}$, the time at which $\left\{Z_k^N\right\}_k$ reaches $N$, since $T_N^*$ coincides with the size of the maximal independent set constructed by this algorithm.

We use the notation in the work of Feng and Kurtz (2006) for the discrete time Markov processes case. Let $\tilde{Y}^N = \left\{\tilde{Y}_k^N\right\}_{k \geq 0}$ be a scaled version of the described process: $\tilde{Y}_k^N = \frac{Z_k^N}{N}$. The transition operator of the process $\tilde{Y}^N$ for $x \in E^N = \left\{\frac{k}{N} : k = 0, 1, ..., N\right\}$ is:

$$T_N(f)(x) := T_{\tilde{Y}^N}(f)(x) = \mathbb{E}\left[f\left(x + \frac{1}{N} + \frac{1}{N}\zeta_{N,x}\right)\right], \tag{2.1}$$

where $\zeta_{N,x}$ is the number of unexplored neighbors of the selected active vertex given that there are already $Nx$ explored vertices. Then $\zeta_{N,x}$ has a Binomial distribution with parameters $N - Nx - 1$ and $\frac{c}{N}$. We consider the embedding maps $\eta_N : E^N \to E$, where $E = [0, 1]$. Define the following continuous process:

$$Y_t^N = \tilde{Y}_{[Nt]}^N = \frac{Z_{[Nt]}}{N} \text{ if } t \in [0, 1]. \tag{2.2}$$

This process is a semimartingale; moreover, it can be decomposed as

$$Y_t^N = \int_0^t \left[1 + c\left(1 - Y_s^N - \frac{1}{N}\right)\right] \mathrm{d}s + \frac{M_{tN}^N}{N},$$

where $\left\{M_t^N\right\}_t$ is a $\mathbb{F}^N = \left\{\mathcal{F}_t^N\right\}_t$ martingale with $\mathcal{F}_t^N = \sigma\left(Z_{[Ns]}^N : 0 \leq s \leq t\right)$.

In Bermolen et al. (2017a) it is proved that the sequence of processes $\left\{Y^N\right\}_N$, contained in the space of càdlàg functions $D_E[0, 1]$, converges in the Skorohod topology to $\{z(t) \wedge 1\}_{0 \leq t \leq 1}$, where $z$ is the solution of the ODE:

$$\dot{z} = 1 + c(1 - z); \quad z(0) = 0. \tag{2.3}$$

This equation has an explicit solution given by $z(t) = \frac{1+c}{c}(1 - e^{-ct})$. Moreover, a law of large numbers can be deduced for the proportion of vertices that form the independent set constructed by the algorithm. In particular, it is proved that $\frac{T_N^*}{N}$ converges in probability to $T^*$ defined by $z(T^*) = 1$, i.e. $T^* = \frac{1}{c}\log(1+c)$.

In the same paper (Bermolen et al., 2017a) and for a different scaling of the process, a diffusion result is also proved from which a central limit theorem for $\frac{T_N^*}{N}$ is deduced: $\sqrt{N}\left(\frac{T_N^*}{N} - T^*\right)$ converges in distribution to a centered normal random variable with variance $\sigma^2 = \frac{c}{2(c+1)^2}$. Now in the present document, we study an LDP for both the sequence of processes $\left\{Y^N\right\}_N$ and for $\left\{\frac{T_N^*}{N}\right\}_N$. It is known that the results of the central limit theorems and large deviations types are independent of each other, and neither is stronger than the other. However, we will see that an LDP also automatically provides results of the law of large numbers type.

2.2. *Large Deviation Principle.* This paper aims to present a more refined analysis of the simple exploration algorithm presented in the previous section. As a corollary, in the next section, we deduce an LDP for the sequence of random variables $\left\{\frac{T_N^*}{N}\right\}_N$.

**Theorem 2.1** (LDP for $\left\{Y^N\right\}_N$)**.** *The sequence $\left\{Y^N\right\}_N$ with $Y^N = \left\{Y_t^N\right\}_{0\le t\le 1}$, where $Y_t^N = \frac{Z^N_{[Nt]}}{N}$, verifies an LDP on $D_E[0,1]$ with good rate function $I : D_E[0,1] \to [0,+\infty]$ such that:*

$$I(\varphi) = \begin{cases} \int_0^1 L(\varphi, \dot\varphi)\, dt & \text{if } \varphi \in \mathcal{H}_L, \\ +\infty & \text{in other case,} \end{cases} \tag{2.4}$$

*where $E = [0,1]$, $L : E \times \mathbb{R} \to \mathbb{R}$ is the cost function given by*

$$L(x,\beta) = \begin{cases} (\beta-1)\left[\log\left(\frac{\beta-1}{c(1-x)}\right) - 1\right] + c(1-x), & \text{if } x < 1 \text{ and } \beta > 1, \\ c(1-x), & \text{if } x < 1 \text{ and } \beta = 1, \\ 0, & \text{if } x = 1 \text{ and } \beta = 0, \\ +\infty & \text{in other cases ,} \end{cases} \tag{2.5}$$

*and $\mathcal{H}_L$ is the set of all absolutely continuous[2] function $\varphi : [0,1] \to [0,1]$ with value 0 at 0 and such that the integral $\int_0^1 L(\varphi(t), \dot\varphi(t))\, dt$ exists and it is finite.*

The proof is deferred to Section 3.

*Remark* 2.2 (Law of large numbers). The cost function (2.5) is the Legendre transform w.r.t the second variable of the function $H\colon E \times \mathbb{R} \to \mathbb{R}$ given by

$$H(x,\alpha) = \begin{cases} \alpha + c(1-x)(e^\alpha - 1), & \text{if } 0 \le x < 1, \\ 0, & \text{if } x = 1, \end{cases} \tag{2.6}$$

that is $L(x,\beta) = \sup_{\alpha\in\mathbb{R}}\{\alpha\beta - H(x,\alpha)\}$. Since $H(x,\alpha)$ is convex with respect to $\alpha$, the function $L$ is also convex with respect to $\beta$ and verifies $H(x,\alpha) = \sup_{\beta\in\mathbb{R}}\{\alpha\beta -$

[2] A function $\varphi : [a,b] \to \mathbb{R}$ is absolutely continuous if it can be written as an integral function; i.e. there exists a Lebesgue integrable function $\psi$ on $[a,b]$ such that $\varphi(x) = \varphi(a) + \int_a^x \psi(t)dt$ for all $x \in [a,b]$.

$L(x,\beta)\}$. We use the notation $H \leftrightarrow L$ for short. As $L(x,\beta) = 0$ if and only if $\beta = H_\alpha(x,0)$, where $H_\alpha(x,\alpha)$ is the partial derivative of $H(x,\alpha)$ w.r.t. $\alpha$, the trajectories with zero cost are the ones that verify $\dot{\varphi} = H_\alpha(\varphi,0) = 1 + c(1-\varphi(t))$. For the initial condition $\varphi(0) = 0$, as expected, the unique trajectory that has zero cost is the fluid limit $z$ given by Equation (2.3) i.e. $I(z) = 0$ and $I(\varphi) > 0$ for all $\varphi \neq z$.

The following proposition gives an intuitive interpretation of the cost function $L(x,\beta)$ in terms of the rate function for the average of independent Poisson random variables.

**Proposition 2.3.** *For $x < 1$ and $\beta > 1$, it is verified that $L(x,\beta) = \Lambda^*_{c(1-x)}(\beta - 1)$, where $\Lambda^*_\lambda(u)$ is the LD rate function for the average of independent Poisson random variables with parameter $\lambda$.*

*Proof*: The rate function given by Crámer's theorem for the average of independent random variables Poisson with parameter $\lambda$ is $\Lambda^*_\lambda(u) = u\left(\log\left(\frac{u}{\lambda}\right) - 1\right) + \lambda$ (see Dembo and Zeitouni (1998) for example). To complete the proof it is enough to observe that $L(x,\beta)$ coincides with $\Lambda^*_\lambda(u)$ when $\lambda = c(1-x)$ and $u = \beta - 1$. □

The previous result can be explained using the following heuristics (which, of course, are far from a proof but give some intuition):

- The graph's sparsity implies that the graph is locally tree-like and that the exploration does not see neighbors of a given vertex being neighbors between them.
- The asymptotic distribution of the number of unexplored neighbors of the selected active vertex is Poisson with a time-varying mean. In other words, the exploration does not change the Poisson nature of the degree distribution, which can be explained by the fact that the biased size distribution of Poisson distribution is again Poisson.

More precisely, the cost of a given curve $x(t)$ such that $x \in \mathcal{H}_L$ with $\dot{x}(t) > 1$ for all $t \in [0,1]$ is given by $L(x(t),\dot{x}(t)) = \Lambda^*_{\lambda(t)}(\dot{x}(t) - 1)$, with $\lambda(t) = c(1-x(t))$. For a fixed $t \in (0,1)$, the curve $x(t)$ represents the macroscopic proportion of explored vertices at time $t$. Then, the infinitesimal increment $\dot{x}(t) \approx \frac{x(t+h)-x(t)}{h}$ corresponds to the mean number of new explored nodes in one step (the new active node and its unexplored blocked neighbors), that is:

$$\frac{Y^N_{t+h} - Y^N_t}{h} = \frac{1}{Nh}\sum_{k=[Nt]+1}^{[Nt+Nh]}\left(1+\zeta^N_k\right) \approx 1 + \frac{1}{Nh}\sum_{k=[Nt]+1}^{[Nt+Nh]}\zeta^N_k,$$

where $\zeta^N_k$ has a Binomial distribution with parameters $N - Z_k - 1$ and $\frac{c}{N}$. For large values of $N$ and $k \in [[Nt]+1, [Nt+Nh]]$, if $\frac{Z_k}{N}$ is close to $x(t)$, then $\zeta^N_k$ can be approximated by a Poisson random variable with parameter $(N - Z_k - 1)\frac{c}{N} \approx c(1-x(t))$. Observe that, in particular, the mean macroscopic behavior $z(t)$ should verify $\dot{z}(t) = 1 + c(1-z(t))$, which is the fluid limit we have already seen. Moreover, the global cost of a deviation from a trajectory $x(t)$ can be interpreted as a consequence of the accumulated cost of microscopic deviations of the average of Poisson random variables of parameter $c(1-x(t))$.

2.2.1. *Rare event probability estimation.* We now use the previous theorem to estimate probabilities of rare events related to $\{Y^N\}_N$. In the next section, we apply these results to derive an LDP for the size of the independent set constructed by the algorithm.

As a consequence of Theorem 2.1, if $A \subset D_E[0,1]$ is a *good set* for $I$ (or an *I-continuous set*, see (Dembo and Zeitouni, 1998)), then $\lim_N \frac{1}{N} \log \mathbb{P}\left(Y^N \in A\right) = -\inf_{\varphi \in A} I(\varphi)$ . The next proposition will facilitate the computation of this infimum for the sets $A$ of interest.

**Proposition 2.4** (Rate function optimization)**.** (1) *The optimization problem for the rate over a set of trajectories $A \subset D_E[0,1]$ can be reduced to a one-dimensional optimization problem: $\inf_{\varphi \in A} I(\varphi) = \inf_{\{\alpha_0 \in \mathbb{R}:\, \hat{x}_{\alpha_0} \in \bar{A}\}} F(\alpha_0)$, where the closure of $A$ is considered with respect to the Skorohod topology,*

$$F(\alpha_0) = \int_0^{T_{\alpha_0}} L\left(x_{\alpha_0}(t), \dot{x}_{\alpha_0}(t)\right) dt, \tag{2.7}$$

*$x_{\alpha_0}$ is the solution of the ODE:*

$$\begin{cases} \dot{x} = 1 + c(1-x)e^{\alpha}, & x(0) = 0, \\ \dot{\alpha} = c(e^{\alpha} - 1), & \alpha(0) = \alpha_0, \end{cases} \tag{2.8}$$

*$T_{\alpha_0} = \inf\{t \in [0,1] :\, x_{\alpha_0}(t) \geq 1\}$ and $\hat{x}_{\alpha_0}(t) = x_{\alpha_0}(t) \wedge 1$.*

(2) *The explicit solution of Equation (2.8) is the fluid limit (2.3) when $\alpha_0 = 0$. For $\alpha \neq 0$ it is given by:*

$$x_{\alpha_0}(t) = \left[\frac{1}{ck_0} \log\left(\frac{1-k_0}{1-k_0 e^{ct}}\right) + \frac{1}{e^{-ct} - k_0} - \frac{1}{1-k_0}\right]\left(e^{-ct} - k_0\right), \tag{2.9}$$

*where $k_0 = 1 - e^{-\alpha_0}$. In this case, $F(\alpha_0)$ can be written as $F(\alpha_0) = \int_0^{T_{\alpha_0}} c\left(1 - x_{\alpha_0}(t)\right)\left[e^{\alpha(t)}\left(\alpha(t) - 1\right) + 1\right] dt$, where $\alpha(t) = -\log\left(1 - k_0 e^{ct}\right)$.*

Then, in other words, Theorem 2.1 and the previous proposition ensure that, given that the process $Y^N \in A$, one might expect that $\sup_{t \in [0,1]} \left|Y_t^N - \hat{x}_{\alpha_0^*}(t)\right| \approx 0$ for some $\alpha_0^*$ such that $\hat{x}_{\alpha_0^*} \in \bar{A}$.

*Proof*: To prove the first statement, note that if $\varphi \in \mathcal{H}_L$ is such that $\varphi(t) = 1$ for all $t \geq t_0$, then $I(\varphi) = \int_0^1 L(\varphi, \dot{\varphi}) \mathrm{d}t = \int_0^{t_0} L(\varphi, \dot{\varphi}) \mathrm{d}t$, so just consider the Euler-Lagrange (EL) equation (2.10) for $x < 1$ and $\beta > 1$. Equation (2.10) gives conditions for a function $\varphi$ to be a *stationary curve* of the functional $I$:

$$L_x(\varphi, \dot{\varphi}) - \frac{\mathrm{d}}{\mathrm{d}t} L_\beta(\varphi, \dot{\varphi}) = 0 \qquad \text{(Euler-Lagrange)}, \tag{2.10}$$

where $L_x$ and $L_\beta$ are the partial derivatives of $L$ w.r.t. $x$ and $\beta$ respectively. In this case, the path $\{x(t)\}_t$ is a stationary curve of $I$ if it satisfies the following ODE:

$$\begin{cases} (x-1)\ddot{x} + (cx - (1+c))\,\dot{x} - cx + (1+c) = 0, \\ x(0) = 0, \\ \dot{x}(0) = v_0. \end{cases} \tag{2.11}$$

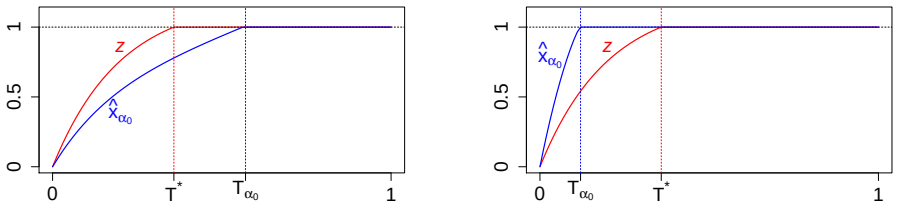


FIGURE 2.1. Graph of $\hat{x}_{\alpha_0}$ for same value of $\alpha_0 < 0$ (left graph) and $\alpha_0 > 0$ (right graph) compared with the fluid limit $z \wedge 1$.

To solve (2.11), we consider Hamilton's equations, which are equivalent to EL (see Arnold (1987), for example):

$$\begin{cases} \dot{x} = H_\alpha(x, \alpha), \\ \dot{\alpha} = -H_x(x, \alpha), \end{cases} \qquad \text{(Hamilton)}, \tag{2.12}$$

where $\alpha$ is an auxiliary function. $H_x$ and $H_\alpha$ are the partial derivatives of $H$ w.r.t. $x$ and $\alpha$. In our case these equations give (2.8). We are interested in solutions $x_{\alpha_0}$ of (2.8) up to the time they reach the value 1, then we take $\hat{x}_{\alpha_0}$ as in the proposition and get $\inf_{\varphi \in A} I(\varphi) = \inf_{\{\alpha_0 : \hat{x}_{\alpha_0} \in \bar{A}\}} I(\hat{x}_{\alpha_0})$.

The uniqueness of the solution of the ODE in (2.11), ensures that a monotony property with respect to the initial condition $\alpha_0$ holds. This implies that $x_{\alpha_0}(t) > t$ for all $t$ if $\alpha_0 > -\infty$, then $T_{\alpha_0} = \inf\{t \in [0,1] : x_{\alpha_0}(t) \geq 1\} \leq 1$ and $I(\hat{x}_{\alpha_0}) = F(\alpha_0)$ with $F(\alpha_0)$ defined in (2.7). Figure 2.1 contains the graph of $\hat{x}_{\alpha_0}$ for same value of $\alpha_0 < 0$ and $\alpha_0 > 0$ compared with the fluid limit $z \wedge 1$.

To prove the second part of the proposition, observe that the fluid limit (2.3) (until it reaches $x = 1$) is a solution of $\dot{z} = 1 + c(1 - z)$, so it is a solution of (2.8) with $\alpha = 0$. If $\alpha_0 \neq 0$, the solution $x_{\alpha_0}$ can be found explicitly and it is given by (2.9). We use that $x_{\alpha_0}$ is solution of (2.8) for the simplification of the cost function $L(x_{\alpha_0}, \dot{x}_{\alpha_0})$. Figure 2.2 contains the graph of $F(\alpha_0) = I(\hat{x}_{\alpha_0})$ as a function of $\alpha_0$. □

*Remark* 2.5. Let us introduce some comments on the previous result:

(1) The ODE continuity theorem is verified with respect to the initial condition for the system (2.8). Then, the solution $\hat{x}_{\alpha_0}$ with initial conditions $x(0) = 0$ and $\alpha(0) = \alpha_0 \approx 0$, is close to the fluid limit $z \wedge 1$.

(2) The system (2.12) is conservative: if $u(t) = (x(t), \alpha(t))$ is the solution of (2.12) with initial conditions $u_0 = (0, \alpha_0)$, it verifies $\dot{u} = J\nabla H(u)$ with $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Since $J$ is an antisymmetric matrix, it results that $\frac{\mathrm{d}}{\mathrm{d}t} H(u) = (\nabla H(u))^t J \nabla H(u) = 0$ for all $t$. Then, the solutions of the general equation (2.12) are contained in the level sets of the Hamiltonian $H$.

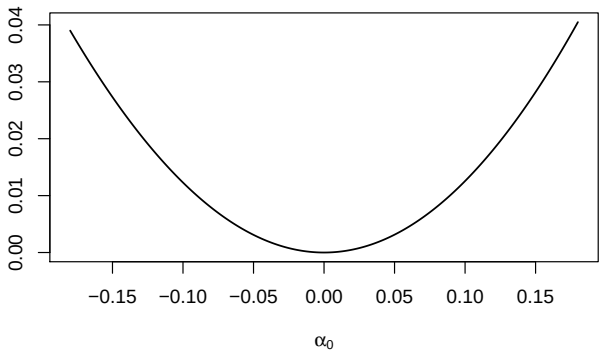


FIGURE 2.2. Graph of the function $F(\alpha_0) = I(\hat{x}_{\alpha_0})$, that is, a parametric version of the rate function.

2.3. *LDP for the size of the independent set constructed by the algorithm.* In the previous section, we presented a path-space LDP for the exploration process defined in Section 1. In this section, we derive from this theorem and the previous proposition about the rate optimization over a specific set, an LDP for the sequence of random variables $\left\{\frac{T_N^*}{N}\right\}_N$. This theorem provides quantitative results for the probability of the independent set's size being bigger/smaller than selected bounds.

**Theorem 2.6.** *Consider $T_N^*$ defined before as the stopping time of the greedy exploration process over $G(N, \frac{c}{N})$.*

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

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

*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*

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

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

*In both cases $F(\alpha_0)$ and $T_{\alpha_0}$ are as in Proposition 2.4.*

*Proof*: We only prove the first statement because the proof of the second one is analogous. Define the set $A_\varepsilon$ such that $A_\varepsilon = \{\varphi \in D_E[0,1] : \varphi(0) = 0, \varphi$ is increasing, $0 \leq \varphi(t) \leq 1$ for all $t$ and $\inf\{t : \varphi(t) = 1\} \geq T^* + \varepsilon\}$. By construction, $A_\varepsilon$ is a good set for $I$, then

$$\lim_N \frac{1}{N} \log \mathbb{P}\left(\frac{T_N^*}{N} \geq T^* + \varepsilon\right) = \lim_N \frac{1}{N} \log \mathbb{P}\left(Y^N \in A_\varepsilon\right) = -\inf_{\{\alpha_0 : \hat{x}_{\alpha_0} \in A_\varepsilon\}} F(\alpha_0).$$

Let $x_{\alpha_0}$ be the solution of the homogenous ODE (2.11) with initial velocity $v_0 = \dot{x}_{\alpha_0}(0) = 1 + ce^{\alpha_0}$. The uniqueness of the solution ensures that the following

monotony property with respect to the initial condition is verified:

$$\text{if } \alpha_0 < \alpha_1 \Rightarrow x_{\alpha_0}(t) < x_{\alpha_1}(t) \text{ for all } t \Rightarrow T_{\alpha_0} > T_{\alpha_1}.$$

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 $x_{\alpha_0^*}(T^* + \varepsilon) = 1$ and

- if $\alpha_0 \leq \alpha_0^* \Rightarrow T_{\alpha_0} \geq T^* + \varepsilon \Rightarrow \hat{x}_{\alpha_0} \in A_\varepsilon$,
- if $\alpha_0 > \alpha_0^* \Rightarrow T_{\alpha_0} < T^* + \varepsilon \Rightarrow \hat{x}_{\alpha_0} \notin A_\varepsilon$,

which implies that $\inf_{\{\alpha_0 : \hat{x}_{\alpha_0} \in A_\varepsilon\}} F(\alpha_0) = \inf_{\{\alpha_0 \leq \alpha_0^*\}} F(\alpha_0)$. To complete the proof it suffices to prove that $\inf_{\{\alpha_0 \leq \alpha_0^*\}} F(\alpha_0) = F(\alpha_0^*)$. Let $h(\alpha_0, t) = L(x_{\alpha_0}, \dot{x}_{\alpha_0})$ and $\alpha_1 < \alpha_2 < 0$. Using the monotony that we mentioned before, it can be seen that $\frac{\partial}{\partial \alpha_0} h(\alpha_0, t) < 0$ for all $\alpha_0 < 0$ and $t \in [0, 1]$, that is $h(\alpha_1, t) > h(\alpha_2, t)$ for all $t$. Finally, since $T_{\alpha_1} > T_{\alpha_2}$ we obtain:

$$F(\alpha_2) = \int_0^{T_{\alpha_2}} L(x_{\alpha_2}, \dot{x}_{\alpha_2}) \mathrm{d}t \leq \int_0^{T_{\alpha_2}} L(x_{\alpha_1}, \dot{x}_{\alpha_1}) \mathrm{d}t < \int_0^{T_{\alpha_1}} L(x_{\alpha_1}, \dot{x}_{\alpha_1}) \mathrm{d}t = F(\alpha_1),$$

which completes the proof. □

2.4. *On the size of the maximum independent set.* The problem of finding the *maximum* independent sets in deterministic graphs is known to be NP-hard. An interesting research question is to find classes on random graphs where finding *maximum* independent sets can be (at least at the first order in $N$) obtained with polynomial complexity. This question is, of course, an instance of a more general viewpoint which aims at identifying phase transitions in the analysis of combinatorial optimization problems, allowing to describe drastically different scenarios depending on a few macroscopic parameters, sometimes called order-parameters.

This type of results has been proven to hold for Erdös-Rényi graphs and configuration models in Spitzer (1975) and Jonckheere and Saenz (2019). The order-parameter being $c$ the mean number of neighbors of a given node. Interestingly the phase transition does not correspond for the graph to be subcritical ($c < 1$) but to a much finer property of the landscape of maximal independent sets. The phase transition corresponds to $c < e$ and differentiates between regimes where a simple degree-greedy algorithm reaches (asymptotically) the maximum independent set or not. This same phase transition is reflected in the properties of the spectrum of the graph, see Coste and Salez (2018).

We conjecture that the large deviations characteristics of the greedy algorithm for discovering maximum independent set also have an interesting transition for values of $c$ around $e$. Since the exact optimal order-one asymptotic value of the maximal independent set's size is known only for values of $c < e$, we cannot yet display a full characterization of this phenomena. We can, however, obtain interesting numerical results by using the Erdös bound, instead of the true value. Let $\sigma_N$ the *maximum* size of the independent set of an ER graph $G(N, c/N)$, then a.s. $\sigma_N \leq \frac{2\log(c)}{c} N(1 + o(N))$ if $c \geq 3$. In Figure 2.3 we compute the large deviation rate corresponding to the event $\{\frac{T_N^*}{N} \geq \sigma_i^*(c)\}$ for $i = 1, 2$. Here $\sigma_1^*$ is the exact proportion of the maximum independent set of an ER graph $G(N, c/N)$ when $c < e$ (Jonckheere and Saenz (2019)) and it is given by $\sigma_1^*(c) = w(c) + \frac{c}{2}(w(c))^2$ with

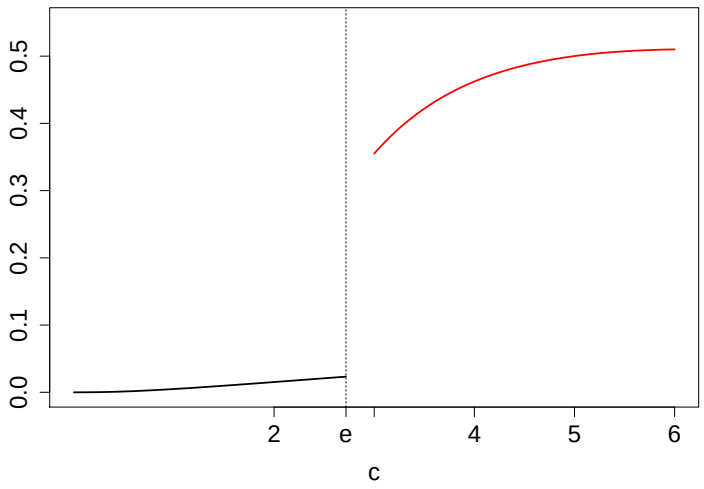


FIGURE 2.3. Evolution of $F(\alpha_0(\sigma_1^*(c)))$ for $0 < c < e$ and $F(\alpha_0(\sigma_2^*(c)))$ for $c \geq 3$.

$w(c) = e^{-W(c)}$ and $W(x)$ the Lambert function. The value $\sigma_2^*(c) = \frac{2}{c}\log(c)$ is the Erdös upper bound for the proportion of the maximum independent set for $c \geq 3$.

Though the numerical computations for $c > e$ could give largely overestimated values, we believe it nevertheless illustrates the clear change of regime around the value $e$. It shows that the independent sets geometry changes, leading to significantly greater large deviations constants for the greedy exploration when $c$ gets larger than $e$. This characterization of the "energy" landscape is a usual situation in statistical physics where interesting phase transitions can be well described through large deviations, see Touchette (2009).

## 3. Proof of Theorem 2.1

In this section, we first briefly describe the theory and main results of Feng and Kurtz (2006) in our context, and then we prove that the previously defined sequence of processes $\{Y_N\}_N$ verifies their assumptions. We organize the main assumptions in four steps described below.

3.1. *Theory of Feng and Kurtz in our context.* As mentioned in Section 1, Feng and Kurtz based their study of large deviations on the *exponential tightness* method. The main result according to this approach is Bryc's theorem. This theorem states that if $(\mathcal{X}, d)$ is a Polish space; $\{\mathbb{P}_N\}_N$ is an exponentially tight [3] sequence of

[3]The sequence $\{\mathbb{P}_N\}_N$ is exponentially tight if for all $\alpha > 0$ exists a compact $K_\alpha \subset \mathcal{X}$ such that $\limsup_N \frac{1}{N}\log \mathbb{P}_N\left(K_\alpha^c\right) \leq -\alpha$.

probability measures defined on $\mathcal{X}$, and the following limit exists:

$$\Lambda(f) = \lim_N \frac{1}{N} \log \int_{\mathcal{X}} e^{Nf(x)} \mathrm{d}\mathbb{P}_N(x) \quad \forall f \in C_b(\mathcal{X}),$$

then $\{\mathbb{P}_N\}_N$ satisfies an LDP with rate function $I : \mathcal{X} \to \mathbb{R}$ such that $I(x) = \sup\{f(x) - \Lambda(f) : f \in C_b(\mathcal{X})\}$, where $C_b(\mathcal{X})$ is the space of bounded and continuous functions $f : \mathcal{X} \to \mathbb{R}$.

Consider now the case in which $\{\mathbb{P}_N\}_N$ comes from a sequence of continuous or discrete-time Markov processes $\left\{Y^N\right\}_N$ with space of states $E_N$. Suppose that $E_N \subset E$ for all $N$. Then $\mathcal{X} = D_E[0,T]$ ($T \leq +\infty$), the space of càdlàg functions equipped with the Skorohod topology (in the discrete case, the time is transformed to be continuous, as we did for the process $\left\{Z_k^N\right\}_k$ in (2.2)). There are results in the literature that ensure equivalent conditions to the exponential tightness, but the calculation of $\Lambda(f)$ is very difficult or even impossible. The theory of Feng and Kurtz solves both, this problem and the exponential tightness. As the transitions characterize the Markov dynamics, instead of calculating $\lim_N \frac{1}{N} \log \mathbb{E}\left[e^{Nf\left(Y^N\right)}\right]$, the convergence of the *Fleming semigroups* (see Fleming (1985)) is studied: $V_t^N : Dom\left(V_t^N\right) \subset B(E) \to B(E)$ such that $V_t^N(f)(x) = \frac{1}{N} \log \mathbb{E}\left[e^{Nf(Y_t^N)} | Y_0^N = x\right]$, where $B(E)$ is the space of bounded, Borel measurable functions (i.e. $t$ is fixed and the domain of the functions $f$ is $E$ instead of the much more complex space $D_E[0,T]$). In Feng and Kurtz (2006) it is proved that, under certain assumptions, the convergence of the Fleming semigroups ensures an LDP. Actually, instead of studying the convergence of $\left\{V_t^N\right\}_{t\in[0,T]}$, the convergence of their (nonlinear) generators $H^N$ is studied. The general idea is that if there is a functional $\mathbf{H}$ such that $H_N \to \mathbf{H}$ (the type of convergence will depend on each case), $\mathbf{H}$ generates a semigroup $\mathbf{V} = \{V_t\}_t$ and the *exponential compact containment condition* is verified, then the sequence $\left\{Y^N\right\}_N$ verifies an LDP with rate function $I$ that depends on $\mathbf{V}$. Moreover, if $\mathbf{H}$ is such that $\mathbf{H}(f)(x) = H(x, f'(x))$ for all $f \in C^1(E)$ and Conditions **8.9**, **8.10** and **8.11** of Feng and Kurtz (2006) are also verified, we obtain a variational version of $I$. In our particular case, the rate will be written as an action integral of $L(x,\beta) \leftrightarrow H(x,\alpha)$.

The main steps of the proof of Theorem 2.1 are now briefly outlined. As a consequence of the first two steps, the process $\{Y^N\}_N$ 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 $T^N$ the transition operators defined in Equation 2.1 and $H_N(f) = \log\left(e^{-Nf} T^N\left(e^{Nf}\right)\right)$.

**S**tep 1. *Verify the convergence of the sequence of operators $H_N$ and derive the limit operator $\mathbf{H}$.* See Proposition 3.2 and note that $\mathbf{H}(f)(x) = H(x, f'(x))$ for $f \in C^1(E)$.

**S**tep 2. *Verify the exponential compact containment condition.* The sequence $\left\{Y^N\right\}_N$ verifies the *exponential compact containment condition* if for all $\alpha > 0$, there exists $K_\alpha \subset E$ compact such that $\limsup_N \frac{1}{N} \log \mathbb{P}\left(\left\{\exists t \in [0,T] : Y_t^N \notin K_\alpha\right\}\right) \leq -\alpha$. In our case $E$ is compact, so this condition is trivially verified by taking $K_\alpha = E$.

**S**tep 3. *Prove that* $\mathbf{H}$ *generates a semigroup* $\mathbf{V} = \{V_t\}_t$ (comparison principle). This is the most technical step. By definition, $V_t^N$ verifies $\frac{\mathrm{d}}{\mathrm{d}t}V_t^N(f) = H_N\left(V_t^N(f)\right)$. Then $\mathbf{H}$ generates a semigroup if there is $\mathbf{V} = \{V_t\}_t$ such that for all $f \in Dom(\mathbf{V})$,

$$\frac{\mathrm{d}}{\mathrm{d}t}V_t(f) = \mathbf{H}\left(V_t(f)\right); \quad V_0(f) = f.$$

The theorem of Crandall and Liggett (1971) implies that $\mu_N(t) = \left(Id - \frac{t}{N}\mathbf{H}\right)^{-N}$ converges to the solution of the previous equation if $\mathbf{H}$ is m-dissipative. Then we need to prove that for all $h \in C(E)$ and $\beta > 0$, there exists $f \in C^1(E)$ such that

$$f - \beta\mathbf{H}(f) - h = 0. \tag{3.1}$$

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* (see Definition 3.3) for Equation (3.1) 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 Feng and Kurtz (2006)). As mentioned by Feng and Kurtz (2006), the verification of the comparison principle is an analytic issue and often gives the impression of being rather involved and disconnected from the probabilistic large deviations problems. An in-depth study of the comparison principle for Hamilton-Jacobi equations in this context is presented in Kraaij (2016), using results from Crandall et al. (1992) and Chapter 9 of Feng and Kurtz (2006). We follow these ideas to prove the comparison principle in our case. See Proposition 3.4.

Once we have verified these three steps, Theorem **6.14** from Feng and Kurtz (2006) assures that $\left\{Y^N\right\}_N$ is exponentially tight and satisfies an LDP with rate function $I$ defined implicitly in terms of $V_t$. This is a theoretical result but does not provide a useful characterization of the rate. The next step provides a simplified version of the rate that can be used in practice.

**S**tep 4. *Construct a variational representation for the rate function* $I$. Let $L(x,\beta) \leftrightarrow H(x,\alpha)$. We state the following result:

**Theorem 3.1.** *If Conditions* ***8.9***, ***8.10*** *and* ***8.11*** *of Feng and Kurtz (2006) are also verified, then:*

(a) $V_t(f) = \mathcal{V}_t(f)$ *for all* $f \in Dom(\mathbf{V})$, *where*

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

*is the Nisio semigroup (see Nisio (1976, 1978); Fleming (1999); El Karoui et al. (1982)) associated to the cost function* $-L$. $\mathcal{Y}$ *is a control subset that we define in subsection 3.4.*

(b) $I(\mathbf{x}) = \inf_{\{\lambda:\,(\mathbf{x},\lambda)\in\mathcal{Y}\}} \left\{ \iint_{U\times[0,1]} L(\mathbf{x}(s),u)\lambda(du\times ds) \right\}$.

(c) *Moreover, the rate function can be written as an action integral:*

$$I(\mathbf{x}) = \int_0^1 L(\mathbf{x}(s),\dot{\mathbf{x}}(s))ds,$$

*if* $\mathbf{x} \in \mathcal{H}_L$ *and* $I((x)) = +\infty$ *in another case.*

*Proof*: The first two sentences, (a) and (b), are proved by Theorems **8.14**, **8.23**, **8.27** and **8.29** of Feng and Kurtz (2006) taking $E = [0,1]$, $U = \mathbb{R}$, the linear operator $A : C^1(E) \to M(E \times U)$ such that $A(f)(x,u) = f'(x)u$, $L(x,\beta) \leftrightarrow H(x,\alpha)$ and $\Gamma = E \times U$. For (c) we use that $L$ is convex w.r.t. the second variable and Jensen's inequality. □

Then, it remains for us to verify Conditions **8.9**, **8.10** and **8.11**. We do this in Section 3.4.

We organize the proof of Theorem 2.1 using the steps mentioned above, that are presented as propositions. As mentioned before, Step 2 is trivially verified in our case.

3.2. *Step 1: Convergence of the nonlinear operators.* Let $H_N : Dom(H_N) \subset B(E) \to B(E)$ such that $H_N = \log\left[e^{-Nf(x)} T_N\left(e^{Nf}\right)(x)\right]$ with $T_N$ the transition operator for the process $\left\{\frac{Z_k^N}{N}\right\}_k$.

**Proposition 3.2.** *There exists a functional* $\mathbf{H}$ *such that* $H_N$ *converges to* $\mathbf{H}$ *when* $N \to \infty$ *in the following sense:* $\lim_{N\to\infty} \sup_{x\in E^N} |H_N(f)(x) - \mathbf{H}(f)(x)| = 0$ *for all* $f \in C^1(E)$. *The functional* $\mathbf{H} : C^1(E) \to B(E)$ *is such that* $\mathbf{H}(f)(x) = H(x, f'(x))$, *where* $H : E \times \mathbb{R} \to \mathbb{R}$ *is defined by*

$$H(x,\alpha) = \begin{cases} \alpha + c(1-x)\left(e^{\alpha} - 1\right), & \text{if } x < 1, \\ 0, & \text{if } x = 1. \end{cases} \tag{3.2}$$

*Proof*: Let us first consider the case where $f \in C^2(E)$. Let $x \in E^N$, $x \neq 1$, and $\zeta_{N,x}$ be the number of unexplored neighbors of the selected vertex, given that there are already $Nx$ explored vertices, then

$$H^N(f)(x) = \begin{cases} \log \mathbb{E}\left[\exp\left\{\frac{f\left(x+\frac{1}{N}+\frac{\zeta_{N,x}}{N}\right)-f(x)}{\frac{1}{N}}\right\}\right], & \text{if } 0 \leq x < 1, \\ 0, & \text{if } x = 1. \end{cases}$$

It is enough to prove that

$$\lim_{N\to\infty} \sup_{x\in E^N\setminus\{1\}} \mathbb{E}\left[\exp\left\{\frac{f\left(x+\frac{1}{N}+\frac{\zeta_{N,x}}{N}\right)-f(x)}{\frac{1}{N}}\right\}\right] - \mathbb{E}\left[e^{f'(x)(1+\zeta_{N,x})}\right] = 0,$$

and this is verified since both $\mathbb{E}\left[e^{f'(x)(\zeta_{N,x}+1)}\left(e^{\pm\frac{M_f}{2}\frac{(\zeta_{N,x}+1)^2}{N}} - 1\right)\right]$ converge to zero, being $M_f = \sup_{\theta\in[0,1]} |f''(\theta)| < \infty$. If $x = 1$, then $H^N(f)(1) = \mathbf{H}(f)(1) = 0$ for all $N$. The result can be extended for $f \in C^1(E)$ by taking a sequence $\{f_m\}_m \subset C^2(E)$ such that $\lim_{m\to\infty} \sup_{x\in E} |f_m(x) - f(x)| = 0$ and the triangular inequality. □

3.3. *Step 3: Comparison principle.* As mentioned before, the verification that for all $\beta > 0$ and $h \in C(E)$ there exists a solution $f \in C^1(E)$ for the equation $f(x) - \beta H(x, f'(x)) - h(x) = 0$ is difficult or imposible. An alternative is to prove the existence (and uniqueness) of viscosity solutions. Moreover, due to Theorem **6.14** of Feng and Kurtz (2006), it is enough to prove that the *comparison principle*

is verified for this Hamilton-Jacobi equation. The ideas to prove it were taken from Kraaij (2016), Chapter 9 of Feng and Kurtz (2006) and Crandall et al. (1992).

Let $\beta > 0$, $h \in C(E)$ and $F_{\beta,h} : E \times \mathbb{R}^2 \to \mathbb{R}$ such that $F_{\beta,h}(x, \epsilon, p) = \epsilon - \beta H(x, p) - h(x)$. Consider the following Hamilton-Jacobi equation:

$$F_{\beta,h}(x, f(x), f'(x)) = 0 \ \ \forall x \in E. \tag{3.3}$$

**Definition 3.3.** The function $\mu \to \mathbb{R}$ is a (viscosity) subsolution [*supersolution*] of Equation (3.3) if it is bounded, upper [*lower*] semi-continuous (u.s.c.) [*l.s.c*] and for all $\phi \in C^1(E)$ and $x_0 \in E$ such that $\mu - \phi$ has a maximum [*minimum*] at $x_0$, we have $F_{\beta,h}(x_0, \mu(x_0), \phi'(x_0)) \leq 0$ [$\geq 0$]. Equation (3.3) verifies the *comparison principle* if for any subsolution $\mu$ and supersolution $v$, it is verified that $\mu \leq v$.

If the comparison principle is verified, then if there is a viscosity solution (both sub and supersolution), it is unique. In Chapter 9 of Feng and Kurtz (2006) algorithms are suggested for constructing sequences $x_\alpha$, $y_\alpha$ (with $\alpha \to +\infty$) such that $(x_\alpha, y_\alpha) \to (z, z)$ and $z$ verifies $\mu(z) - v(z) = \sup_{x \in E} \{\mu(x) - v(x)\}$.

**Proposition 3.4.** *For each $\beta > 0$ and $h \in C(E)$ the comparison principle is satisfied for Equation 3.3 with $f \in C^1(E) = Dom(\boldsymbol{H})$.*

*Proof*: Let $\mu$ be a subsolution and $v$ a supersolution of Equation (3.3).
Let $\psi : [0,1]^2 \to \mathbb{R}^+$ such that $\psi(x, y) = \frac{1}{2}(x - y)^2$ and let $x_\alpha, y_\alpha \in E$ such that

$$\mu(x_\alpha) - v(y_\alpha) - \alpha\psi(x_\alpha, y_\alpha) = \sup_{x,y \in E} \{\mu(x) - v(y) - \alpha\psi(x, y)\}.$$

As consequence of Proposition **4.2** in Kraaij (2016) it is enough to prove that the following inequality holds:

$$\liminf_{\alpha \to \infty} H(x_\alpha, \alpha\psi_x(x_\alpha, y_\alpha)) - H(y_\alpha, \alpha\psi_x(x_\alpha, y_\alpha)) \leq 0,$$

where $\psi_x$ is the derivative of $\psi$ w.r.t. $x$. If $z \in [0, 1)$, then

$$H(x_\alpha, \alpha\psi_x(x_\alpha, y_\alpha)) - H(y_\alpha, \alpha\psi_x(x_\alpha, y_\alpha)) = -c\left(e^{\alpha(x_\alpha - y_\alpha)} - 1\right)(x_\alpha - y_\alpha).$$

By Proposition **3.7** in Crandall et al. (1992), we know that $x_\alpha - y_\alpha \to 0$ and due to Lemma 4.5 in Kraaij (2016) we have:

$$\sup_\alpha H(y_\alpha, \alpha(x_\alpha - y_\alpha)) = \sup_\alpha \alpha(x_\alpha - y_\alpha) + c(1 - y_\alpha)\left(e^{\alpha(x_\alpha - y_\alpha)} - 1\right) < \infty,$$

and this implies that $\sup_\alpha \alpha(x_\alpha - y_\alpha) < \infty$. Then $\{\alpha(x_\alpha - y_\alpha)\}_\alpha$ has a convergent subsequence. Let $A$ be its limit. Then,

$$\liminf_{\alpha \to \infty} H(x_\alpha, \alpha\psi_x(x_\alpha, y_\alpha)) - H(y_\alpha, \alpha\psi_x(x_\alpha, y_\alpha)) \leq H(z, A) - H(z, A) = 0.$$

For $z = 1$, we repeat the previous analysis, being careful with cases in which $x_\alpha = 1$ or $y_\alpha = 1$ after a certain $\alpha_0$. □

3.4. *Step 4: Variational representation of the rate function.* In this section, we formally define the Nisio semigroup $\mathcal{V}_t$ that was mentioned in Theorem 3.1. By Theorem 3.1, it is enough to prove that Conditions **8.9**, **8.10**, and **8.11** from Feng and Kurtz (2006) are verified in our case. We present them as propositions. Als, the role of absolutely continuous functions in the definition of the rate function $I$ is explained.

**Definition 3.5** (Control set of a linear operator and Nisio semigroup)**.** Let $U$ and $E$ be complete and separable metric spaces. Let $A : Dom(A) \subset B(E) \to M(E \times U)$ be a single valued linear operator. Let $\mathcal{M}_m(U)$ be the space of Borel measures $\lambda$ on $U \times [0,1]$ satisfying $\lambda(U \times [0,t]) = t$ for all $t \in [0,1]$. The measure $\lambda$ is known as a *relaxed control*. We say that the pair $(\mathbf{x}, \lambda) \in D_E[0,1] \times \mathcal{M}_m(U)$ satisfies the *relaxed control equation* for $A$ if and only if:

(1) $\iint_{U\times[0,t]} |A(f)(\mathbf{x}(s), u)| \, \lambda(\mathrm{d}u \times \mathrm{d}s) < \infty \; \forall f \in Dom(A), \; \forall t \in [0,1]$;

(2) $f(\mathbf{x}(t)) - f(\mathbf{x}(0)) = \iint_{U\times[0,t]} A(f)(\mathbf{x}(s), u) \lambda(\mathrm{d}u \times \mathrm{d}s) \; \forall f \in Dom(A), \; \forall t \in [0,1]$.

We denote the collection of pairs satisfying the above properties by $\mathcal{Y}$. The Nisio semigroup corresponding to the control problem determined by the linear operator $A$ and the cost function $-L$ is:

$$\mathcal{V}_t(f)(x_0) = \sup_{\{(\mathbf{x},\lambda)\in\mathcal{Y}:\, \mathbf{x}(0)=x_0\}} \left\{ f(\mathbf{x}(t)) - \iint_{U\times[0,t]} L(\mathbf{x}(s), u) \lambda(\mathrm{d}u \times \mathrm{d}s) \right\} \tag{3.4}$$

for each $x_0 \in E$ (the supremum of an empty set is defined to be $-\infty$). Note that operator $A$ appears in the definition of the control set.

In our case, as $\mathbf{H}(f)(x) = H(x, f'(x))$ for each $x \in E = [0,1]$ and $H \leftrightarrow L$, 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 $U = \mathbb{R}$ and $A : C^1(E) \to M(E \times U)$ is the linear operator $A(f)(x,u) = f'(x)u$. As $L$ is convex w.r.t. $\beta$, it follows that a deterministic control $\lambda(\mathrm{d}u \times \mathrm{d}s) = \delta_{u(s)}(\mathrm{d}u)\mathrm{d}s$ is allways the control with smallest cost by Jensen's inequality. Moreover, if $\mathbf{x} : E \to \mathbb{R}$ is an absolutely continuous function (we note $\mathbf{x} \in \mathcal{AC}$ for short), then

$$f(\mathbf{x}(t)) - f(\mathbf{x}(0)) = \int_0^t f'(\mathbf{x}(s)) \, \dot{\mathbf{x}}(s) \mathrm{d}s = \iint_{\mathbb{R}\times[0,t]} f'(\mathbf{x}(s)) \; u\lambda(\mathrm{d}u \times \mathrm{d}s),$$

if we define $\lambda = \lambda(\mathbf{x})$ such that $\lambda(\mathrm{d}u \times \mathrm{d}s) = \delta_{\dot{\mathbf{x}}(s)}(\mathrm{d}u)\mathrm{d}s$. Then, the supremum in Equation 3.4 is reached on $\{(\mathbf{x}, \lambda) : \mathbf{x} \in \mathcal{AC}, \; \mathbf{x}(0) = x_0\} \subset \mathcal{Y}$.

**Proposition 3.6.** *Conditions* **8.9** *of Feng and Kurtz (2006) are verified.*

*Proof*: Conditions (1) to (4) are trivially verified. For (5) we construct $\Psi_f$ as in Lemma **10.21** of Feng and Kurtz (2006). □

**Proposition 3.7.** *Condition* ***8.10*** *from Feng and Kurtz (2006) is verified, i.e. for all $x_0 \in E$ there exists $(\mathbf{x}, \lambda) \in \mathcal{Y}$ such that $\mathbf{x}(0) = x_0$ and*

$$\iint_{U\times[0,1]} L(\mathbf{x}(s), u) \, \lambda(du \times ds) = 0.$$

*Proof*: Since $L(x,u) = 0 \Leftrightarrow u = H_\alpha(x,0)$, the function $q(x) = H_\alpha(x,0) = 1 + c(1-x)$ solves the equation $L(x,q(x)) = 0$ for all $x \in E$. Note that the fluid limit verifies $\dot{z} = q(z)$ with the initial condition $z(0) = 0$. Given $x_0 \in E$, there exists $t_0 \in [0,1]$ such that $z(t_0) = x_0$, then $z(t+t_0)$ is the solution of $\dot{x} = q(x)$ with $x(0) = x_0$. Define $\mathbf{x}(t) = z(t+t_0) \wedge 1$ for all $t \in [0,1]$ and $\lambda$ such that $\lambda(\mathrm{d}u \times \mathrm{d}s) = \delta_{\{q(\mathbf{x}(s))\}}(\mathrm{d}u) \times \mathrm{d}s$, then $(\mathbf{x}, \lambda) \in \mathcal{Y}$ and verifies the required condition. □

**Proposition 3.8.** *Condition* **8.11** *from Feng and Kurtz (2006) is verified, i.e. for all $x_0 \in E$ and $f \in C^1(E)$ there exists $(\mathbf{x}, \lambda) = (\mathbf{x}_f, \lambda_f) \in \mathcal{Y}$ such that $\mathbf{x}(0) = x_0$ and*

$$\int_{t_1}^{t_2} \mathbf{H}(f)(\mathbf{x}(s))\, ds \leq \iint_{U \times [t_1,t_2]} [A(f)(\mathbf{x}(s),u) - L(\mathbf{x}(s),u)]\, \lambda(ds \times du),$$

*for all $0 \leq t_1 \leq t_2 \leq 1$.*

*Proof*: Let $x_0 \in E$ and $f \in C^1(E)$ be fixed. Since $\mathbf{H}(f)(x) \geq A(f)(x,u) - L(x,u)$ for all $(x,u) \in E \times U$, we need to find $(\mathbf{x}, \lambda) \in \mathcal{Y}$ such that $\mathbf{x}(0) = x_0$ and

$$\int_0^t H(\mathbf{x}(s), f'(\mathbf{x}(s)))\, \mathrm{d}s = \iint_{U \times [0,t]} [f'(\mathbf{x}(s))\, u - L(\mathbf{x}(s),u)]\, \lambda(\mathrm{d}s \times \mathrm{d}u), \quad (3.5)$$

for all $t \in [0,1]$. If we define $q_f(x) = H_\alpha(x, f'(x))$, then $H(x, f'(x)) = f'(x) q_f(x) - L(x, q_f(x))$ and Equation 3.5 is verified if we take $\lambda(\mathrm{d}u \times \mathrm{d}s) = \delta_{\{q_f(\mathbf{x}(s))\}}(\mathrm{d}u)\ \mathrm{d}s$. Now we have to add conditions so that in addition $(\mathbf{x}, \lambda) \in \mathcal{Y}$. In particular, $(\mathbf{x}, \lambda)$ has to verify:

$$\int_0^t g'(\mathbf{x}(s))\, q_f(\mathbf{x}(s))\, \mathrm{d}s = g(\mathbf{x}(t)) - g(\mathbf{x}(0)) \ \ \forall t \in [0,1],\ \forall g \in C^1(E).$$

Then we look for a path that solves the following problem:

$$\begin{cases} \mathbf{x} \text{ is differentiable almost everywhere and } \dot{\mathbf{x}}(t) = q_f(\mathbf{x}(t)), \\ \mathbf{x}(0) = x_0, \\ \mathbf{x}(t) \in [0,1] \text{ for all } t \geq 0. \end{cases} \quad (3.6)$$

Let $x_0 \in [0,1)$. Note that $q_f(x) = 1 + c(1-x)e^{f'(x)} > 1$ is continuous, then from Peano's theorem (see Crandall (1972)) we know that the ODE $\begin{cases} \dot{\mathbf{x}}(t) = q_f(\mathbf{x}(t)) \\ \mathbf{x}(t_0) = x_0 \end{cases}$ has a local solution $\mathbf{x} : J \to \mathbb{R}$, being $J$ an open neighborhood of $t_0$, it is also increasing and verifies $\mathbf{x}(t) \geq t$ for all $t \in [0,1]$. Since we need $\mathbf{x} \in D_E[0,1]$, we can paste these local solutions until the time $T_{x_0}$ it reaches 1 and define $\mathbf{x}(t) = 1$ for $T_{x_0} \leq t \leq 1$. If $x_0 = 1$, we take $\mathbf{x} \equiv 1$. □

## Acknowledgements

The authors would like to thank the referees for their valuable comments that help us to improve our manuscript significantly.

## References


A. de Acosta. Exponential tightness and projective systems in large deviation theory. In *Festschrift for Lucien Le Cam*, pages 143–156. Springer, New York (1997). doi:10.1007/978-1-4612-1880-7_9. URL https://doi.org/10.1007/978-1-4612-1880-7_9.

B. I. Arnold. *Métodos Matemáticos da Mecanica Clásica.* Editora Mir Moscovo (1987).

F. Baccelli and N. Tien Viet. Generating Functionals of Random Packing Point Processes: From Hard-Core to Carrier Sensing. *Computing Research Repository (CoRR).* (2012). abs/1202.0225.

P. Bermolen, M. Jonckheere and Sanders J. Scaling Limits and Generic Bounds for Exploration Processes. *Journal of Statistical Physics.* **5**, 989–1018 (2017a).

P. Bermolen, M. Jonckheere and Pascal. Moyal. The jamming constant of uniform random graphs. *Stochastic Processes and their Applications.* **7**, 2138–2178 (2017b).

Simon Coste and Justin Salez. Emergence of extended states at zero in the spectrum of sparse random graphs (2018). 1809.07587.

H. Cramér. Sur un nouveau théoreme-limite de la theorie des probabilitiés. *Acta. Sci. et Ind.* **736**, 5–23 (1938).

M. G. Crandall and T. M. Liggett. Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.* **93**, 265–298 (1971). ISSN 0002-9327. doi:10.2307/2373376. URL https://doi.org/10.2307/2373376.

Michael G. Crandall. A generalization of Peano's existence theorem and flow invariance. *Proc. Amer. Math. Soc.* **36**, 151–155 (1972). ISSN 0002-9939. doi:10.2307/2039051. URL https://doi.org/10.2307/2039051.

Michael G. Crandall, Hitoshi Ishii and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* **27** (1), 1–67 (1992). ISSN 0273-0979. doi:10.1090/S0273-0979-1992-00266-5. URL https://doi.org/10.1090/S0273-0979-1992-00266-5.

Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition (1998). ISBN 0-387-98406-2. doi:10.1007/978-1-4612-5320-4. URL https://doi.org/10.1007/978-1-4612-5320-4.

M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I, II (and III). *Comm. Pure Appl. Math.* **28**, 1–47; ibid. 28 (1975), 279–301 (1975). ISSN 0010-3640. doi:10.1002/cpa.3160280102. URL https://doi.org/10.1002/cpa.3160280102.

Paul Dupuis and Richard S. Ellis. *A weak convergence approach to the theory of large deviations.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York (1997). ISBN 0-471-07672-4. doi:10.1002/9781118165904. URL https://doi.org/10.1002/9781118165904. A Wiley-Interscience Publication.

Paul Dupuis, Richard S. Ellis and Alan Weiss. Large deviations for Markov processes with discontinuous statistics. I. General upper bounds. *Ann. Probab.* **19** (3), 1280–1297 (1991). ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(199107)19:3<1280:LDFMPW>2.0.CO;2-D&origin=MSN.

N. El Karoui, J.P. Lepeltier and B. Marchal. Nisio semi-group associated to the control of Markov processes. *Stochastic Differential Systems.* **43** (1982).

J.W. Evans. Random and cooperative sequential adsorption. *Reviews of Modern Physics.* **65** (4), 1281–1329 (1993). .

L. C. Evans and H. Ishii. A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2** (1), 1–20 (1985). ISSN 0294-1449. URL http://www.numdam.org/item?id=AIHPC_1985__2_1_1_0.

Jin Feng and Thomas G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2006). ISBN 978-0-8218-4145-7; 0-8218-4145-9. doi:10.1090/surv/131. URL https://doi.org/10.1090/surv/131.

Pablo A. Ferrari, Roberto Fernández and Nancy L. Garcia. Perfect simulation for interacting point processes, loss networks and Ising models. *Stochastic Process. Appl.* **102** (1), 63–88 (2002). ISSN 0304-4149. doi:10.1016/S0304-4149(02)00180-1. URL https://doi.org/10.1016/S0304-4149(02)00180-1.

Wendell H. Fleming. Exit probabilities and optimal stochastic control. *Appl. Math. Optim.* **4** (4), 329–346 (1977/78). ISSN 0095-4616. doi:10.1007/BF01442148. URL https://doi.org/10.1007/BF01442148.

Wendell H. Fleming. A stochastic control approach to some large deviations problems. In *Recent mathematical methods in dynamic programming (Rome, 1984)*, volume 1119 of *Lecture Notes in Math.*, pages 52–66. Springer, Berlin (1985). doi:10.1007/BFb0074780. URL https://doi.org/10.1007/BFb0074780.

W.H. Fleming. Stochastic analysis, control optimization and applications. (1999). A volume in honor of W.H. Fleming.

M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York (1984). ISBN 0-387-90858-7. doi:10.1007/978-1-4684-0176-9. URL https://doi.org/10.1007/978-1-4684-0176-9. Translated from the Russian by Joseph Szücs.

David Gamarnik and Madhu Sudan. Limits of local algorithms over sparse random graphs. *Ann. Probab.* **45** (4), 2353–2376 (2017). ISSN 0091-1798. doi:10.1214/16-AOP1114. URL https://doi.org/10.1214/16-AOP1114.

Hamed Hatami, László Lovász and Balázs Szegedy. Limits of locally-globally convergent graph sequences. *Geom. Funct. Anal.* **24** (1), 269–296 (2014). ISSN 1016-443X. doi:10.1007/s00039-014-0258-7. URL https://doi.org/10.1007/s00039-014-0258-7.

M. Jonckheere and M. Saenz. Asymptotic optimality of degree-greedy discovering of independent sets in configuration model graphs (2019). URL https://arxiv.org/abs/1808.10358.

Dieter Jungnickel. *Chapter 5, The Greedy Algorithm, In:Graph, Networks and Algorithm*, volume 5 of *Algorithm and Computation in Mathematics*. Springer-Verlag Berlin Heidelberg, second edition (2005). URL https://doi.org/10.1007/b138283.

R. M. Karp and M. Sipser. Maximum matching in sparse random graphs. In *22nd Annual Symposium on Foundations of Computer Science (sfcs 1981)*, pages 364–375 (1981). doi:10.1109/SFCS.1981.21.

L. Kleinrock and H. Takagi. Throughput analysis for persistent csma systems. *IEEE Transactions on Communications* **33** (7), 627–638 (1985).

Richard Kraaij. 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** (2), 321–345 (2016). ISSN 0022-4715. doi:10.1007/s10955-016-1542-8. URL https://doi.org/10.1007/s10955-016-1542-8.

Colin McDiarmid. Colourings of random graphs. In *Graph colourings (Milton Keynes, 1988)*, volume 218 of *Pitman Res. Notes Math. Ser.*, pages 79–86. Longman Sci. Tech., Harlow (1990).

M. Nisio. On stochastic optimal controls and envelope of Markovian semi-groups. *Proc. of. Intern. Symp. SDE* pages 297–325 (1976).

M. Nisio. On non linear semi-groups for Markov processes associated with optimal stopping. *Applied Math. and Opt.* **4**, 143–169 (1978).

George L. O'Brien and Wim Vervaat. Compactness in the theory of large deviations. *Stochastic Process. Appl.* **57** (1), 1–10 (1995). ISSN 0304-4149. doi:10.1016/0304-4149(95)00007-T. URL https://doi.org/10.1016/0304-4149(95)00007-T.

Mathew D. Penrose. Random parking, sequential adsorption, and the jamming limit. *Comm. Math. Phys.* **218** (1), 153–176 (2001). ISSN 0010-3616. doi:10.1007/s002200100387. URL https://doi.org/10.1007/s002200100387.

B. Pittel. On the probable behaviour of some algorithms for finding the stability number of a graph. *Mathematical Proceedings of the Cambridge Philosophical Society* **92** (3), 511–526 (1982). doi:10.1017/S0305004100060205.

A. Puhalskii. The method of stochastic exponentials for large deviations. *Stochastic Process. Appl.* **54** (1), 45–70 (1994). ISSN 0304-4149. doi:10.1016/0304-4149(94)00004-2. URL https://doi.org/10.1016/0304-4149(94)00004-2.

T. Ritchie. Construction of the thermodynamic jamming limit for the parking process and other exclusion schemes on $Z^d$. *Journal of Statistical Physics* **122**, 381–398 (2006).

Frank Spitzer. Markov random fields on an infinite tree. *Ann. Probability* **3** (3), 387–398 (1975). ISSN 0091-1798. doi:10.1214/aop/1176996347. URL https://doi.org/10.1214/aop/1176996347.

Hugo Touchette. The large deviation approach to statistical mechanics. *Phys. Rep.* **478** (1-3), 1–69 (2009). ISSN 0370-1573. doi:10.1016/j.physrep.2009.05.002. URL https://doi.org/10.1016/j.physrep.2009.05.002.

Nicholas C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** (4), 1217–1235 (1995). ISSN 1050-5164. URL http://links.jstor.org/sici?sici=1050-5164(199511)5:4<1217:DEFRPA>2.0.CO;2-A&origin=MSN.