The Annals of Applied Probability
2000, Vol. 10, No. 1, 104–122

LARGE DEVIATIONS AT EQUILIBRIUM FOR A LARGE STAR-SHAPED LOSS NETWORK

BY CARL GRAHAM AND NEIL O'CONNELL

CMAP, École Polytechnique and BRIMS, Hewlett-Packard Labs

We consider a symmetric network composed of N links, each with capacity C. Calls arrive according to a Poisson process, and each call concerns L distinct links chosen uniformly at random. If each of these links has free capacity, the call is held for an exponential time; otherwise it is lost. The semiexplicit stationary distribution for this process is similar to a Gibbs measure: it involves a normalizing factor, the partition function, which is very difficult to evaluate. We let N go to infinity and keep fixed the rate of call attempts concerning any link. We use asymptotic combinatorics and recent techniques involving the law of large numbers to obtain the asymptotic equivalent for the logarithm of the partition function and then the large deviation principle for the empirical measure of the occupancies of the links. We give an explicit formula for the rate function and examine its properties.

1. Introduction. We consider a large network composed of links numbered 1 to N, each with capacity C. Calls arrive as a Poisson process. Each call chooses a route, uniformly at random, in the set of all possible routes,

$$\mathscr{R}^N = \{\text{subsets of } L \text{ distinct links among } 1, 2, \dots, N\}. \tag{1.1}$$

The call is lost if any link on the chosen route is full, otherwise it holds one channel on each of these links for an exponential time, after which it releases all of these L channels simultaneously. Arrivals, route choices, and call durations are all independent.

This model corresponds to many situations of simultaneous service, such as telecommunication or computer networks, locking of items in data-bases, parallel computing or job processing in factories. It has historically been called "star-shaped," since in the case $L = 2$ it corresponds to a communication network with N terminal nodes joined by raylike links to a central station, through which calls are routed. See [13], [14], [9] and [8].

A (huge) Markovian description of the network is given by $\mathbf{Y}^N = (Y_r^N)_{r\in\mathscr{R}^N}$, where the process Y_r^N counts the number of ongoing calls on route r. The process

$$X_i^N = \sum_{r\in\mathscr{R}^N:\ i\in r} Y_r^N \tag{1.2}$$

counts the number of occupied channels on link i, and the simultaneous releases prevent $(X_i^N)_{1\le i\le N}$ from being Markovian. The processes Y_r^N and

Received November 1998; revised May 1999.

AMS 1991 *subject classifications*. Primary 60F10, 60K35, 60G10; secondary 68M10, 90B12.

Key words and phrases. Large networks, stationary distribution, partition function, large deviations.

X_i^N have sample paths in the Skorohod space $\mathbb{D}(\mathbb{R}_+, \{0, 1, \dots, C\})$. A relevant tractable quantity is the empirical measure $\mu^N = (1/N)\sum_{i=1}^N \delta_{X_i^N}$, which has samples in the space $\mathscr{P}(\mathbb{D}(\mathbb{R}_+, \{0, 1, \dots, C\}))$ of probability measures on the Skorohod space, and its process of time-marginals,

$$\bar{X}^N = (\bar{X}_t^N)_{t\geq 0}, \qquad \bar{X}_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_i^N(t)}. \tag{1.3}$$

We can identify $\mathscr{P} = \mathscr{P}(\{0, 1, \dots, C\})$ with the C-dimensional simplex, and $\bar{X}^N$ has sample paths in $\mathbb{D}(\mathbb{R}_+, \mathscr{P})$ and records the occupancies of the links averaged over the network.

In this paper, we obtain a large deviation principle (LDP) for the stationary distribution of $\bar{X}^N$, as N goes to infinity. We keep the arrival rate seen by each link equal to ν, thus the arrival rate on any route r in $\mathscr{R}^N$ is given by

$$\nu_N = \frac{\nu}{\binom{N-1}{L-1}}. \tag{1.4}$$

An accepted call lasts an exponential time of parameter λ.

We may note that the evolution of a link j is given by the action of the generator of the Markov process $\mathbf{Y}^N$ on functions depending only on X_j^N. In this asymptotic regime, the contribution of call arrivals can be expressed asymptotically as a fixed function of X_j^N and of the empirical measure of L-tuples of the $(X_i^N)_{1\leq i\leq N}$, which corresponds to relatively weak L-body mean-field interaction in statistical mechanics terminology. The simultaneous release of L channels at each call termination, along the corresponding route, yields much more complex terms, corresponding to more local and strong interaction. See [4], [5] and [6].

We shall not directly use such dynamical expressions. We start from a classical semiexplicit expression for the stationary distribution, given, for example, in [13] and [14]. It contains a normalizing term, the partition function, which must be evaluated. We express this term as a sum, over the state space of $\bar{X}^N$, of combinatorial terms. Using as a key ingredient the law of large numbers (LLN) obtained [13] for $\bar{X}^N$, we prove that the asymptotic equivalent of the logarithm of this sum can be deduced from the equivalents of the logarithm of its summands near the LLN limit.

The large deviation principle will also require finding the equivalents for the logarithms of the summands at each point of the state space, with some uniformity. We obtain these equivalents using combinatorics related to the occupation problem in [7]. We eventually obtain a LDP with an explicit formula for the rate function. We prove that the rate function has a unique local minimum (which is at the LLN limit), is strictly positive everywhere except at this minimum (at which it vanishes), and is locally uniformly convex at this point. The rate function is not convex in general, which denotes rather strong interaction. We give a fairly complete description of its curvature.

The idea of using laws of large numbers to obtain large deviation principles seems to be quite useful, and has been exploited to an even greater extent than here, for a variety of applications, by O'Connell [10]–[12].

Large deviations results have been obtained for a class of networks, which includes this model, under a different asymptotic regime, by Chang and Wang [1]. In this regime, the topology of the network remains fixed while the capacity and arrival rate go to infinity.

The outline of the paper is as follows. In Section 2, we present some preliminary material. In Section 3, we simplify the asymptotic evaluation of the partition function using Whitt's law of large numbers. In Section 4, we compute the necessary combinatorial asymptotics. We present the large deviation principle in Section 5 and a detailed analysis of the rate function in Section 6. In Section 7 we present some pictures.

2. Preliminaries.

2.1. *Some pathwise results.* Whitt [13] gives a functional law of large numbers (LLN) on $\bar{X}^N$, given that $\bar{X}_0^N$ satisfies a LLN; he deduces a LLN for the stationary distribution. The limit process $(Q_t)_{t\geq 0}$ satisfies the nonlinear ordinary differential equation (ODE),

$$\begin{aligned}
\dot{Q}_t\{0\} &= -\nu(1-Q_t\{C\})^{L-1}Q_t\{0\} + \lambda Q_t\{1\},\\
&\cdots\\
\dot{Q}_t\{k\} &= \nu(1-Q_t\{C\})^{L-1}(Q_t\{k-1\}-Q_t\{k\})\\
&\quad +\lambda((k+1)Q_t(k+1)-kQ_t\{k\}),\\
&\cdots\\
\dot{Q}_t\{C\} &= \nu(1-Q_t\{C\})^{L-1}Q_t(C-1)-\lambda C Q_t\{C\}
\end{aligned} \tag{2.1}$$

on $\mathscr{P}$. We set $\rho = \nu/\lambda$ and $\rho_N = \nu_N/\lambda$. Any fixed point q_ρ of this ODE satisfies

$$kq_\rho\{k\} = \rho\big(1-q_\rho\{C\}\big)^{L-1} q_\rho\{k-1\}, \qquad k = 1, 2, \dots, C, \tag{2.2}$$

which is solved in terms of $q_\rho\{0\}$ as

$$q_\rho\{k\} = q_\rho\{0\}\frac{\Big(\rho\big(1-q_\rho\{C\}\big)^{L-1}\Big)^k}{k!}, \qquad k = 0, 1, \dots, C, \tag{2.3}$$

and such a q_ρ is in $\mathscr{P}$ if and only if we have $q_\rho\{k\} \geq 0$ and

$$\begin{aligned}
&\sum_{k=0}^{C} q_\rho\{k\} = 1 \text{ or equivalently,}\\
&q_\rho\{0\} = \left(\sum_{k=0}^{C}\frac{\Big(\rho\big(1-q_\rho\{C\}\big)^{L-1}\Big)^k}{k!}\right)^{-1}.
\end{aligned} \tag{2.4}$$

Thus q_ρ is a fixed point in $\mathscr{P}$ if and only if the blocking probability $q_\rho\{C\}$ satisfies

$$q_\rho\{C\} = \left(\sum_{k=0}^{C} \frac{\left(\rho(1-q_\rho\{C\})^{L-1}\right)^k}{k!}\right)^{-1} \frac{\left(\rho(1-q_\rho\{C\})^{L-1}\right)^C}{C!} \tag{2.5}$$

in [0, 1]. This corresponds to the Erlang fixed point approximation; see [9] and [14]. The r.h.s. of (2.5) is a continuous decreasing function of $q_\rho\{C\}$ which is strictly positive for $q_\rho\{C\} = 0$ and tends to 0 as $q_\rho\{C\}$ tends to 1, hence there is a unique solution $q_\rho\{C\}$ which is in]0, 1[. Thus there is a unique fixed point q_ρ to (2.1), determined by (2.3) and (2.5), and q_ρ belongs to the interior $\mathscr{P}^o$ of $\mathscr{P}$.

Let the first moment or mean occupancy of $\alpha \in \mathscr{P}$ be $\langle\alpha\rangle = \sum_{k=0}^{C} k\alpha\{k\}$. By (2.2),

$$\langle q_\rho\rangle = \sum_{k=0}^{C} k q_\rho\{k\} = \rho(1-q_\rho\{C\})^L. \tag{2.6}$$

A natural important problem is to obtain rates of convergence for this LLN. Let us recall some results obtained by sample path considerations. Graham and Méléard [3] prove propagation of chaos in total variation norm on path space for an initially empty network: there is a law Q on $\mathbb{D}([0,T],\{0,1,\dots,C\})$, defined by a tree construction and unique solution to a nonlinear martingale problem, such that $|\mathscr{L}(X_1^N,\dots,X_k^N) - Q^{\otimes k}| \le k^2C(T)/N$. This implies the convergence in probability of μ^N to Q and of $\bar{X}^N = (\bar{X}_t^N)_{t\ge 0}$ to $(Q_t)_{t\ge 0}$ and can be extended to more general initial conditions satisfying a LLN. Graham and Méléard [4] and Hunt [8] prove a functional central limit theorem for $\bar{X}^N$ for appropriate initial conditions, with an Ornstein–Uhlenbeck process as the limit. Graham and Méléard [5] and [6] also prove functional large deviation results, complete only for $C = 1$.

2.2. *The stationary distribution.* The irreducible Markov process $\mathbf{Y}^N = (Y_r^N)_{r\in\mathscr{R}^N}$ has a unique stationary distribution on

$$\mathscr{A}^N = \left\{\mathbf{m} = (m_r)_{r\in\mathscr{R}^N} \colon m_r \in \mathbb{N},\ \sum_{r\in\mathscr{R}^N \colon\ i\in r} m_r \le C,\ \forall i \in \{1,\dots,N\}\right\}$$

given using the notation $|\mathbf{m}| = \sum_{r\in\mathscr{R}^N} m_r$ by

$$\begin{aligned}\hat{\pi}^N(\mathbf{m}) = P_{st}(\mathbf{Y}_t^N = \mathbf{m}) &= \frac{1}{Z_\rho^N} \prod_{r\in\mathscr{R}^N} \frac{(\rho_N)^{m_r}}{m_r!}\\ &= \frac{1}{Z_\rho^N} \frac{\rho^{|\mathbf{m}|}}{\binom{N-1}{L-1}^{|\mathbf{m}|}} \prod_{r\in\mathscr{R}^N} \frac{1}{m_r!}, \qquad \mathbf{m} \in \mathscr{A}^N,\end{aligned} \tag{2.7}$$

where the normalizing factor or partition function Z_ρ^N is given by

$$Z_\rho^N = \sum_{\mathbf{m}\in\mathscr{A}^N} \frac{\rho^{|\mathbf{m}|}}{\binom{N-1}{L-1}^{|\mathbf{m}|}} \prod_{r\in\mathscr{R}^N} \frac{1}{m_r!}. \tag{2.8}$$

Computing this factor is a NP-complete problem, and good approximations are needed; for further discussions and references, see [14].

This gives the distribution π^N on $\mathscr{P}$ of $\bar{X}_t^N$ in equilibrium. There is a well-defined function f^N: $\mathscr{A}^N \mapsto \mathscr{P}$ such that $\bar{X}_t^N = f^N(\mathbf{Y}_t^N)$; see (1.2) and (1.3). We set

$$\mathscr{A}^N(\alpha) = \left\{\mathbf{m} \in \mathscr{A}^N\colon f^N(\mathbf{m}) = \alpha\right\}, \qquad \alpha \in \mathscr{P}$$

and obtain

$$\begin{aligned}\pi^N(\alpha) &= P_{st}(\bar{X}_t^N = \alpha)\\ &= \frac{1}{Z_\rho^N} \frac{\rho^{N\langle\alpha\rangle/L}}{\binom{N-1}{L-1}^{N\langle\alpha\rangle/L}} \sum_{\mathbf{m}\in\mathscr{A}^N(\alpha)} \prod_{r\in\mathscr{R}^N} \frac{1}{m_r!}, \qquad \alpha \in \mathscr{P},\end{aligned} \tag{2.9}$$

where we use that (considering the total number of occupied links in the network)

$$\forall\, \mathbf{m} \in \mathscr{A}^N(\alpha), \qquad N\langle\alpha\rangle = L|\mathbf{m}| \le NC. \tag{2.10}$$

3. The law of large numbers and the partition function. Pathwise results do not extend directly to the stationary distributions: it is difficult to exchange the $N \to \infty$ and $t \to \infty$ limits. Whitt [13] achieves this by a compactness–uniqueness method (see Theorem 3 and Section III in [13] or Section 4 in [9]) and obtains the functional law of large numbers below. The LLN result for fixed time was later proved from direct computations on the partition function; see in [14] the extensions in Section 4 of Theorem 2.4 and Corollary 2.6.

THEOREM 3.1. *We have equivalently*, *with q_ρ the fixed point given by* (2.3) *and* (2.5),

$$\lim_{N\to\infty} \pi^N = \delta_{q_\rho} \text{ weakly}, \qquad \lim_{N\to\infty} \bar{X}_t^N = q_\rho \text{ in probability at equilibrium.}$$

This convergence is uniform in t on bounded intervals.

We now investigate the asymptotics of $\log Z_\rho^N$ for large N. We set

$$\sigma(N, \alpha) = \sum_{\mathbf{m}\in\mathscr{A}^N(\alpha)} \prod_{r\in\mathscr{R}^N} \frac{1}{m_r!} \tag{3.1}$$

and for any set B included in $\mathscr{P}$, we obtain from (2.9),

$$\log \pi^N(B) = \log \sum_{\alpha \in B} \frac{\rho^{N\langle\alpha\rangle/L}}{\binom{N-1}{L-1}^{N\langle\alpha\rangle/L}} \sigma(N,\alpha) - \log Z_\rho^N. \tag{3.2}$$

The sum has a support of cardinality less than $(N+1)^C$, since it is included in the set of α in $\mathscr{P}$ such that $N\alpha(0)$, $N\alpha(1), \dots, N\alpha(C)$ are integers. We bound this sum between its maximal term and its maximal term multiplied by $(N+1)^C$ and set

$$K(N,\alpha) = \langle\alpha\rangle \frac{L-1}{L} \log N - \frac{\langle\alpha\rangle}{L} \log(\rho(L-1)!) - \frac{1}{N} \log \sigma(N,\alpha), \tag{3.3}$$

with $K(N,\alpha) = +\infty$ for any α not in the support of the sum, and obtain

$$\log \pi^N(B) = -N \inf_{\alpha \in B} K(N,\alpha) - \log Z_\rho^N + O(\log N). \tag{3.4}$$

We use this formula and the LLN to simplify the evaluation of Z_ρ^N.

THEOREM 3.2. *For any neighborhood B of the law q_ρ given by* (2.3) *and* (2.5),

$$\log Z_\rho^N = -N \inf_{\alpha \in B} K(N,\alpha) + O(\log N) + o_B(1).$$

PROOF. Theorem 3.1 implies that for any neighborhood B of q_ρ, the left-hand side of (3.4) goes to 0 when N goes to infinity. Hence the result. □

In the future, we shall obtain an asymptotic equivalent for $K(N,\alpha)$, for α in a neighborhood of q_ρ, which is continuous and nonzero, and then use this theorem, for a sequence of open balls B shrinking to q_ρ, in order to deduce an equivalent for $\log Z_\rho^N$.

4. Some asymptotic combinatorics. We restrict our attention at first to α in the state space $\mathscr{P}^N$ for $\bar{X}^N$, given by

$$\mathscr{P}^N = \{\alpha \in \mathscr{P} \colon N\alpha(0), N\alpha(1), \dots, N\alpha(C), N\langle\alpha\rangle/L \in \mathbb{N}\}.$$

LEMMA 4.1. *Let $d(N,\alpha)$ be the number of ways of setting up $N\langle\alpha\rangle/L$ distinguishable calls so that the resulting network occupancy $\mathbf{m}$ is in $\mathscr{A}^N(\alpha)$ (with $|\mathbf{m}| = N\langle\alpha\rangle/L$). Then*

$$d(N,\alpha) = \sum_{\mathbf{m} \in \mathscr{A}^N(\alpha)} \frac{|\mathbf{m}|!}{\prod_{r \in \mathscr{R}^N} m_r!} = |\mathbf{m}|!\sigma(N,\alpha).$$

PROOF. For each $\mathbf{m} \in \mathscr{A}^N(\alpha)$, the multinomial coefficient $|\mathbf{m}|!(\prod_{r\in\mathscr{R}^N} m_r!)^{-1}$ counts the number of ways of partitioning $|\mathbf{m}|$ distinguishable balls in successive subsets of size m_r for r in $\mathscr{R}^N$ (which we order for this purpose). Each possible set-up of $|\mathbf{m}|$ calls giving occupancy $\mathbf{m}$ is clearly in one-to-one correspondence with such a partitioning. □

PROPOSITION 4.2. *Let $w(N, \alpha)$ be the number of ways of dropping $|\mathbf{m}| = N\langle\alpha\rangle/L$ distinguishable groups of L distinguishable balls in N distinguishable boxes, so that the balls in each group fall in distinct boxes and that there are $N\alpha\{k\}$ boxes with k balls, $k = 0, 1, \dots, C$. Then $w(N, \alpha) = L!^{|\mathbf{m}|}d(N, \alpha)$ and hence*

$$\sigma(N, \alpha) = w(N, \alpha)\left(\frac{N\langle\alpha\rangle}{L}!\, L!^{N\langle\alpha\rangle/L}\right)^{-1}.$$

PROOF. Each box corresponds to a link, and each group of L balls to a call. There are $L!$ different ways to drop these L balls in a given subset of L boxes (corresponding to a given route), hence $w(N, \alpha) = L!^{|\mathbf{m}|}d(N, \alpha)$. We express $d(N, \alpha)$ using Lemma 4.1 to obtain the formula for $\sigma(N, \alpha)$. □

The definition of $w(N, \alpha)$ recalls the occupancy problem (see [7] II-5) with the additional constraint that balls in a group must not fall in the same box. This corresponds essentially to the fact that routes are constituted of L *distinct* links.

PROPOSITION 4.3. *Let $w_+(N, \alpha)$ be the number of ways of dropping $N\langle\alpha\rangle/L$ distinguishable groups of L distinguishable balls in N distinguishable boxes, so that there are $N\alpha\{k\}$ boxes with k balls, $k = 0, 1, \dots, C$. This is simply the number of ways of dropping $N\langle\alpha\rangle$ balls in N boxes with that given occupancy of the boxes. Given one of these ways of dropping the balls, let $a(N, \alpha)$ be the number of permutations of the balls for which balls in a group do not fall in the same box. Then*

$$w_+(N, \alpha) = \frac{N!}{\prod_{k=0}^{C}(N\alpha\{k\})!}\frac{(N\langle\alpha\rangle)!}{\prod_{k=0}^{C}(k!)^{N\alpha\{k\}}} \geq w(N, \alpha) = w_+(N, \alpha)\frac{a(N, \alpha)}{(N\langle\alpha\rangle)!}.$$

PROOF. This formula expresses $w_+(N, \alpha)$ as the product of two multinomial coefficients, the first counting all possible partitions of the N boxes in subsets of $N\alpha\{k\}$ boxes which are to contain k balls, the second all possible possible partitions of the $N\langle\alpha\rangle$ balls in $N\alpha\{k\}$ subsets of k balls, $k = 0, 1, \dots, C$; see [7] II-5. More precisely, the second multinomial coefficient is the quotient of the number $(N\langle\alpha\rangle)!$ of permutations of the balls in a given configuration counted in $w_+(N, \alpha)$ by the number of such permutations which keep any ball in its original box, and hence give rise to the same global configuration. We can express $w(N, \alpha)$ similarly as $w_+(N, \alpha)$, only replacing $(N\langle\alpha\rangle)!$ by the number $a(N, \alpha)$ of permutations for which balls in a group do not fall in the same box. □

We wish to show that $w(N, \alpha)$ and $w_+(N, \alpha)$ are asymptotically close, with some uniformity on α. We first give a loose lower bound for $w(N, \alpha)$, then an appropriately tight and uniform lower bound on $w(N, \alpha)/w_+(N, \alpha)$.

REMARK. For arbitrarily large N we may find α such that $w(N, \alpha) = 0$ and $w_+(N, \alpha) \geq 1$; for $L = 2$ and $C = 2$ we take $\alpha(2) = 1/N$ and $\alpha(0) = 1 - 1/N$.

LEMMA 4.4. *Let α in $\mathscr{P}^N$ be such that $N\langle\alpha\rangle \geq CL$. Then*

$$w(N,\alpha) \geq \frac{N!}{\prod_{k=0}^{C}(N\alpha\{k\})!} L!^{N\langle\alpha\rangle/L} \geq 1.$$

PROOF. The multinomial number counts the possible choices on the boxes, and $L!^{N\langle\alpha\rangle/L}$ counts the permutations of balls within each group (balls in a group are in distinct boxes, hence these permutations give distinct configurations). Thus, it is sufficient to prove that for $N\langle\alpha\rangle \geq CL$ there is at least one way to place the balls as in the definition of $w(N,\alpha)$, once we have fixed the $N\alpha\{k\}$ boxes which should hold k balls, for $k = 0, 1, \ldots, C$.

We fix these boxes, and call any box which should hold k balls "a box of type k." We now prove by induction on C that for $N\langle\alpha\rangle \geq CL$ there is at least one way to place k balls in every box of type k, so that balls in a group do not fall in the same box. This is obvious for $C = 1$. Let us assume it true for $C - 1 \geq 1$.

For $N\alpha(C) \geq L$, we place the balls by layers. We place balls successively in each box of type 1, then in each box of type 2, and so on until we place balls successively in each box of type C, thus completing the first layer. The boxes of type 1 now hold 1 ball each. We go back and place balls successively in each box of type 2, then in each box of type 3, and so on until we place balls successively in each box of type C, thus completing the second layer. The type 2 boxes now hold 2 balls each. We continue in a similar manner until there are no balls left. Since there are at least L boxes of type C, we never place balls from the same group in the same box, and we eventually fill up all boxes appropriately.

For $N\alpha(C) \leq L-1$, we take a group of L balls and place one ball in each box of type C. There remains $L - N\alpha(C)$ balls in the group to place properly. There is a total of $N\langle\alpha\rangle - N\alpha(C) \geq CL - L + 1$ balls left to place; since the boxes are either of type k with $k \leq C - 1$ or of type C and contain already one ball, there must be at least $(CL - L + 1)/(C - 1) > L$ boxes which can each accept in the future at least one ball. So we can place the remaining $L - N\alpha(C)$ balls of the group in separate boxes distinct from the $N\alpha(C)$ ones already used. Then we are left with $N\langle\alpha\rangle - L \geq (C-1)L$ balls to place according to a configuration in which the maximal number of balls in a box is $C - 1$, and by induction we know there is at least a way of doing so. □

PROPOSITION 4.5. *We have, for any α in $\mathscr{P}^N$,*

$$0 \geq \log\frac{w(N,\alpha)}{w_+(N,\alpha)} \geq O(\log N)$$

with an $O(\log N)$ term uniform for $N \geq 2$ and α in $\mathscr{P}^N$ such that $N\langle\alpha\rangle \geq CL$.

PROOF. The upper bound is obvious (see Proposition 4.3) and $w(N,\delta_0) = w_+(N,\delta_0) = 1$. For $\alpha \neq \delta_0$ we consider N large enough so that $N\langle\alpha\rangle \geq CL$, and shall bound $a(N,\alpha)$ from below. The boxes are fixed, and the order of placement of the balls in the boxes is taken into account: we call "spot" the conjunction of a box and an order of placement in the box. After the $(j-1)$th group of L balls has been placed, $(j-1)L$ spots have been occupied. The first ball in the jth group can thus be placed in at least $N\langle\alpha\rangle - (j-1)L$ spots, the second in $N\langle\alpha\rangle - (j-1)L - C$ spots (the placement of the first ball in a box prevents the placement of the second ball in the at most C spots in the box) and so on, until the last ball in the group can be placed in at least $N\langle\alpha\rangle - (j-1)L - (L-1)C \geq 1$ spots (only the spots in $L-1$ boxes, each with at most C spots, are forbidden). After thus placing groups of L balls for $j = 1, 2, \dots, (N\langle\alpha\rangle/L) - C$, there are C groups left, and Lemma 4.4 applied to this restricted placement problem with fixed boxes states that there is at least $(L!)^C$ ways to do so. Thus

$$a(N,\alpha) \geq (L!)^C \prod_{j=1}^{(N\langle\alpha\rangle/L)-C} \prod_{k=0}^{L-1} (N\langle\alpha\rangle - (j-1)L - kC).$$

Since $(N\langle\alpha\rangle)! = (LC)! \prod_{j=1}^{(N\langle\alpha\rangle/L)-C} \prod_{k=0}^{L-1} (N\langle\alpha\rangle - (j-1)L - k)$ we obtain, using Proposition 4.3,

$$\frac{w(N,\alpha)}{w_+(N,\alpha)} = \frac{a(N,\alpha)}{(N\langle\alpha\rangle)!} \geq \frac{L!^C}{(LC)!} \prod_{j=1}^{(N\langle\alpha\rangle/L)-C} \prod_{k=0}^{L-1} \left(1 - \frac{k(C-1)}{N\langle\alpha\rangle - (j-1)L - k}\right)$$

and a simple bound yields

$$\frac{w(N,\alpha)}{w_+(N,\alpha)} \geq \frac{L!^C}{(LC)!} \prod_{j=1}^{(N\langle\alpha\rangle/L)-C} \left(1 - \frac{(L-1)(C-1)}{N\langle\alpha\rangle - jL + 1}\right)^{L-1}.$$

We take the logarithm. Classically,

$$\begin{aligned}\sum_{j=1}^{(N\langle\alpha\rangle/L)-C} &\log\left(1 - \frac{(L-1)(C-1)}{N\langle\alpha\rangle - jL + 1}\right)\\ &\geq \int_{C+(1/L)-1}^{(N\langle\alpha\rangle/L)+(1/L)-1} \log\left(1 - \frac{(L-1)(C-1)}{xL}\right) dx,\end{aligned}$$

which we compute to obtain

$$\begin{aligned}\log \frac{w(N,\alpha)}{w_+(N,\alpha)} \geq \frac{L-1}{L}\Bigg(&(N\langle\alpha\rangle + 1 - L)\log\left(1 - \frac{(L-1)(C-1)}{N\langle\alpha\rangle + 1 - L}\right)\\ &+(CL+1-L)\log(CL+1-L)\\ &-(L-1)(C-1)\log(N\langle\alpha\rangle + C - CL) - C\log C\Bigg)\\ &+ C\log L! - \log(LC)!\end{aligned}$$

and the right-hand term is $O(\log N)$ with the uniformity we have stated. $\square$

For α and β in $\mathscr{P}$ we define the entropy and relative entropy (or Kullback information),

$$H(\alpha) = -\sum_{k=0}^{C} \alpha\{k\} \log \alpha\{k\}, \qquad H(\alpha|\beta) = \sum_{k=0}^{C} \alpha\{k\} \log \frac{\alpha\{k\}}{\beta\{k\}}, \tag{4.1}$$

(with the conventions $0 \log 0 = 0$, etc.) and the continuous functions

$$\begin{aligned} K(\alpha) = {} & -H(\alpha) + \sum_{k=0}^{C} \alpha\{k\} \log k! \\ & -\langle\alpha\rangle \frac{L-1}{L} (\log\langle\alpha\rangle - 1) - \frac{\langle\alpha\rangle}{L} \log \rho, \end{aligned} \tag{4.2}$$

$$\begin{aligned} J(\alpha) = {} & K(\alpha) - K(q_\rho) = H(\alpha|q_\rho) \\ & - \frac{L-1}{L}\left(\langle q_\rho\rangle - \langle\alpha\rangle + \langle\alpha\rangle \log \frac{\langle\alpha\rangle}{\langle q_\rho\rangle} \right). \end{aligned} \tag{4.3}$$

Sanov's theorem states that $\alpha \mapsto H(\alpha|q_\rho)$ is the rate function for the large deviation principle for the empirical measure of i.i.d. random variables of law q_ρ. We shall prove that the rate function for the LDP for the star-shaped network is given by J. This result can be interpreted as a perturbation of Sanov's theorem, strong enough to modify the convexity properties of the rate function. See [2], Section 2.1.1.

THEOREM 4.6. *We have, for any α in $\mathscr{P}^N$,*

$$K(N, \alpha) = K(\alpha) + O\left(\frac{\log N}{N}\right),$$

see (3.3) *and* (4.2), *with a $O((\log N)/N)$ remainder term bounded below uniformly, and bounded above uniformly for $N \geq 2$ and α in $\mathscr{P}^N$ such that $N\langle\alpha\rangle \geq CL$.*

PROOF. The statement for $\alpha = \delta_0$ is obvious. Otherwise, Propositions 4.2 and 4.3 give

$$\begin{aligned} \log \sigma(N, \alpha) = {} & \log N! - \sum_{k=0}^{C} \log(N\alpha\{k\})! + \log(N\langle\alpha\rangle)! - N \sum_{k=0}^{C} \alpha\{k\} \log k! \\ & - \log \frac{N\langle\alpha\rangle}{L}! - \frac{N\langle\alpha\rangle}{L} \log L! + \log \frac{w(N, \alpha)}{w_+(N, \alpha)}. \end{aligned}$$

Consider the Stirling formula, $\log n! = n(\log n - 1) + O(\log n)$; see [7]. We have $\log 0! = 0(\log 0 - 0)$ and $\log 1! = 1(\log 1 - 1) + 1$. Since the function log is increasing and $\log 2 > 0$, we have $\log NC \geq \log N\langle\alpha\rangle > 0$ for $N\langle\alpha\rangle \geq 2$ and

$\log N \geq \log N\alpha\{k\} > 0$ for $N\alpha\{k\} \geq 2$. We obtain, using these results and Proposition 4.5,

$$\log\,\sigma(N,\alpha) = N(\log\,N-1) - N\sum_{k=0}^{C}\alpha\{k\}(\log\,N\alpha\{k\}-1)$$
$$+ N\langle\alpha\rangle(\log\,N\langle\alpha\rangle - 1) - N\sum_{k=0}^{C}\alpha\{k\}\log\,k!$$
$$-\frac{N\langle\alpha\rangle}{L}\left(\log\frac{N\langle\alpha\rangle}{L}-1\right) - \frac{N\langle\alpha\rangle}{L}\log\,L! + O(\log\,N)$$

with an $O(\log\,N)$ remainder term uniformly bounded above and bounded below uniformly for $N \geq 2$ and α in $\mathscr{P}^N$ such that $N\langle\alpha\rangle \geq CL$. Since $\sum_{k=0}^{C}\alpha\{k\} = 1$, we have

$$\log\,\sigma(N,\alpha) = NH(\alpha) - N\sum_{k=0}^{C}\alpha\{k\}\log\,k! + N\langle\alpha\rangle\frac{L-1}{L}(\log\,N\langle\alpha\rangle - 1)$$
$$-\frac{N\langle\alpha\rangle}{L}\log(L-1)! + O(\log\,N)$$

and we conclude considering (3.3) and (4.2). $\square$

5. The large deviation principle.

THEOREM 5.1. *We have* [*see* (2.3), (2.6) *and* (4.2)]

$$\lim_{N\to\infty}\frac{1}{N}\log Z_\rho^N = -K(q_\rho) = -\log q_\rho\{0\} - \frac{L-1}{L}\langle q_\rho\rangle.$$

PROOF. We use Theorem 3.2 for a sequence of open balls B shrinking to q_ρ and not containing δ_0. We may then use Theorem 4.6 with a uniform remainder. We obtain the result using the continuity of K at q_ρ and expliciting $K(q_\rho)$. $\square$

THEOREM 5.2. *A large deviation principle with continuous rate function J, defined in* (4.3), *holds for $(\pi^N)_{N\geq 1}$; for any Borel set B included in $\mathscr{P}$,*

$$-\inf_{\alpha\in B^o} J(\alpha) \leq \liminf_{N\to\infty}\frac{1}{N}\log\,\pi^N(B) \leq \limsup_{N\to\infty}\frac{1}{N}\log\pi^N(B) \leq -\inf_{\alpha\in B} J(\alpha).$$

We have $J(\alpha) = H(\alpha|q_\rho)$ if and only if $\langle\alpha\rangle = \langle q_\rho\rangle$, else $J(\alpha) < H(\alpha|q_\rho)$. Note that J is continuous in $\mathscr{P}$, C^∞ in its interior $\mathscr{P}^o$ and $J(q_\rho) = 0$.

PROOF. Let a Borel set B be given. We divide equation (3.4) by N and obtain

$$\frac{1}{N}\log\,\pi^N(B) = -\inf_{\alpha\in B}K(N,\alpha) - \frac{1}{N}\log Z_\rho^N + O\left(\frac{\log\,N}{N}\right)$$

and use Theorems 4.6 and 5.1; we recall that $K(N,\alpha) = +\infty$ for $\alpha \notin \mathscr{P}^N$.

The LDP lower bound follows easily from the classical

$$\liminf_{N\to\infty}\Big(-\inf_{\alpha\in B} K(N,\alpha)\Big) = -\limsup_{N\to\infty}\Big(\inf_{\alpha\in B} K(N,\alpha)\Big) \ge -\inf_{\alpha\in B}\Big(\limsup_{N\to\infty} K(N,\alpha)\Big),$$

hence using the continuity of K (we have a limit only for α in $B\cap\mathscr{P}^N$),

$$\liminf_{N\to\infty}\Big(-\inf_{\alpha\in B} K(N,\alpha)\Big) \ge -\inf_{\alpha\in B^o}\Big(\limsup_{N\to\infty} K(N,\alpha)\Big) = -\inf_{\alpha\in B^o} K(\alpha).$$

For the upper bound, we use uniform convergence. Theorem 4.6 yields

$$-\inf_{\alpha\in B} K(N,\alpha) \le -\inf_{\alpha\in B} K(\alpha) + O\Big(\frac{\log N}{N}\Big)$$

with a uniform remainder term, hence

$$\limsup_{N\to\infty}\Big(-\inf_{\alpha\in B} K(N,\alpha)\Big) \le -\inf_{\alpha\in B} K(\alpha)$$

form which we deduce that LDP upper bound.

The last statement follows from the study of the function $y\mapsto 1-y-y\log y$. $\square$

6. The shape of the rate function. So that the LDP may be of any practical use, we give an appropriate description of the rate function J. We already know that it is nonnegative and continuous on $\mathscr{P}$ and C^∞ in $\mathscr{P}^o$, that $J(q_\rho)=0$ and that $J(\alpha)=H(\alpha|q_\rho)$ if and only if $\langle\alpha\rangle=\langle q_\rho\rangle$, otherwise $J(\alpha)<H(\alpha|q_\rho)$.

The function J is defined on $\mathscr{P}$, which is naturally embedded in $\mathbb{R}^{C+1}$. We differentiate twice at α in $\mathscr{P}^o$ considered as a subset of $\mathbb{R}^{C+1}$, but only the restriction of these differentials to $\mathscr{T}_\alpha=\{\beta-\alpha\colon \beta\in\mathscr{P}\}\subset\mathscr{T}=\{h\in\mathbb{R}^{C+1}\colon h_0+h_1+\cdots+h_C=0\}$ is pertinent. The space $\mathscr{T}_\alpha$ can be interpreted as the tangent space of $\mathscr{P}^o$.

Using the last formulation in (4.3), for $i,j=0,1,\dots,C$,

$$\begin{aligned}\partial_i J(\alpha) &= 1+\log\frac{\alpha\{i\}}{q_\rho\{i\}} - i\frac{L-1}{L}\log\frac{\langle\alpha\rangle}{\langle q_\rho\rangle},\\ \partial_{ij}J(\alpha) &= \frac{\delta_{ij}}{\alpha\{i\}} - \frac{L-1}{L}\frac{ij}{\langle\alpha\rangle}.\end{aligned}\tag{6.1}$$

We denote the product of two matrices A and B (the number of columns of A coinciding with the number of rows of B) by AB or $A\cdot B$. If A is a row vector and B a column vector, $A\cdot B$ coincides with the canonical scalar product. We consider $DJ(\alpha)$ as a row vector acting on column vectors h, and the notation $DJ(\alpha)\cdot h$ is coherent with classical differential notation. We have $DJ(q_\rho)=(1,1,\dots,1)$, with $C+1$ terms, and $DJ(q_\rho)\cdot h=0$ for any h in $\mathscr{T}$, which is natural since there is a minimum in $\mathscr{P}$ of J at q_ρ. Somewhat surprisingly, D^2J does not depend on ρ.

THEOREM 6.1. *The law q_ρ is the only point α in $\mathscr{P}^o$ at which $DJ(\alpha)\cdot h = 0$ for all h in $\mathscr{T}$; hence it is the only point in $\mathscr{P}^o$ at which there is a local extremum of J. Moreover q_ρ is the only point in $\mathscr{P}$ at which there is a local minimum, and in particular $J(\alpha) > 0$ for any α in $\mathscr{P} - \{q_\rho\}$, while $J(q_\rho) = 0$.*

PROOF. If $DJ(\alpha)\cdot h = 0$ for all h in $\mathscr{T}$, then $DJ(\alpha)$ is collinear to $(1, \dots, 1)$, and $\partial_i J(\alpha)$ does not depend on i. Hence, for $i = 1, \dots, C$,

$$\log \frac{\alpha\{i\}}{q_\rho\{i\}} = \log \frac{\alpha\{i-1\}}{q_\rho\{i-1\}} + \frac{L-1}{L} \log \frac{\langle\alpha\rangle}{\langle q_\rho\rangle}$$

and using (2.2) and (2.6),

$$\frac{\alpha\{i\}}{\alpha\{i-1\}} = \frac{q_\rho\{i\}}{q_\rho\{i-1\}} \left(\frac{\langle\alpha\rangle}{\langle q_\rho\rangle}\right)^{(L-1)/L} = i^{-1}\langle\alpha\rangle^{(L-1)/L}\rho^{1/L},$$

hence

$$i\alpha\{i\} = \langle\alpha\rangle^{(L-1)/L}\rho^{1/L}\alpha\{i-1\}. \tag{6.2}$$

Summation over i yields

$$\langle\alpha\rangle = \langle\alpha\rangle^{(L-1)/L}\rho^{1/L}(1 - \alpha\{C\}),$$

which implies $\langle\alpha\rangle = \rho(1-\alpha\{C\})^L$, and going back to (6.2), we see that α solves the recurrence relation (2.2) of which we know that the unique solution in $\mathscr{P}$ is q_ρ.

For any boundary point α of $\mathscr{P}$ except δ_0, $\partial_i J(\alpha) = -\infty$ whenever $\alpha\{i\} = 0$, and $\partial_1 J(\delta_0) = -\infty$. Hence there cannot be a local minimum at the boundary. $\square$

We deduce classically exponential estimates, and a strong LLN.

THEOREM 6.2. *There is exponential decay of $\pi^N(A)$ for any closed set A not containing q_ρ, and in equilibrium $(\bar{X}_t^N)_{N\geq 1}$ converges a.s. to q_ρ, for any $t \geq 0$.*

PROOF. Let A be a closed subset of $\mathscr{P}$ not containing q_ρ, which is compact since $\mathscr{P}$ is compact. The continuous function J attains a minimum $m > 0$ on A and, using the LDP upper bound, for any $0 < a < m$, for sufficiently large N, $\pi^N(A) \leq e^{-Na}$. Thus,

$$\sum_{N\geq 1} P_{st}(\bar{X}_t^N \in A) = \sum_{N\geq 1} \pi^N(A) < \infty,$$

hence by the Borel–Cantelli lemma $P_{st}(\bar{X}_t^N \in A$ for infinitely many $N) = 0$. The closed set A being arbitrary, a.s. convergence of $\bar{X}_t^N$ to q_ρ can be deduced easily. $\square$

We denote by A^* the transposed matrix of A. For v the column vector $(0, 1, \dots, C)^*$ and $\operatorname{diag}(\alpha)$ and $\operatorname{diag}(\alpha)^{-1} = \operatorname{diag}(\alpha^{-1})$ the diagonal matrices $\operatorname{diag}(\alpha\{0\}, \alpha\{1\}, \dots, \alpha\{C\})$ and $\operatorname{diag}(\alpha\{0\}^{-1}, \alpha\{1\}^{-1}, \dots, \alpha\{C\}^{-1})$,

$$D^2 J(\alpha) = \operatorname{diag}(\alpha)^{-1} - \frac{L-1}{L}\frac{1}{\langle\alpha\rangle} v v^* \tag{6.3}$$

is the sum of a positive definite matrix of which we know the inverse and of a matrix of rank 1. This particular structure enables us to study its invertibility and signature (number of strictly positive and negative terms in an orthogonal decomposition).

Let $Q(\alpha)$ be the restriction of $D^2 J(\alpha)$ to $\mathscr{T} = \{h\colon h_0 + h_1 + \cdots + h_C = 0\}$, $v^\perp$ be the space of column vectors x orthogonal to v, such that $v^* x = 0$, and $\mathscr{N} = v^\perp \cap \mathscr{T} = \{h \in \mathscr{T}\colon \langle h\rangle = 0\}$. Actually, we consider $\mathscr{N}_\alpha = v^\perp \cap \mathscr{T}_\alpha = \{\beta - \alpha\colon \beta \in \mathscr{P}, \langle\beta\rangle = \langle\alpha\rangle\}$, which has a natural interpretation in terms of mean occupancy.

Denote by $\langle\alpha\rangle_2$ and $\operatorname{Var}(\alpha)$ the second moment and variance of α, respectively. For $\alpha \in \mathscr{P}^o$ we denote by $\bar\alpha$ the probability measure

$$\bar\alpha\{i\} = \frac{i\alpha\{i\}}{\langle\alpha\rangle}, \qquad i = 0, 1, \dots, C. \tag{6.4}$$

Note that $\bar\alpha\{0\} = 0$ and $\bar\alpha\{i\} > 0$ for $i = 1, \dots, C$, and that

$$\begin{aligned} \langle\bar\alpha\rangle &= \frac{\langle\alpha\rangle_2}{\langle\alpha\rangle}, \qquad \langle\bar\alpha - \alpha\rangle = \frac{\operatorname{Var}(\alpha)}{\langle\alpha\rangle} > 0, \\ (\bar\alpha - \alpha)^* \operatorname{diag}(\alpha)^{-1}(\bar\alpha - \alpha) &= \frac{\operatorname{Var}(\alpha)}{\langle\alpha\rangle^2} > 0. \end{aligned} \tag{6.5}$$

For $L \geq 2$ we define on $\mathscr{P}$ and $\mathbb{R}^{C+1}$ the second degree polynomial of several variables

$$F(\alpha) = \frac{L}{L-1}\langle\alpha\rangle - \operatorname{Var}(\alpha) = \langle\alpha\rangle^2 + \Big(1 + \frac{1}{L-1}\Big)\langle\alpha\rangle - \langle\alpha\rangle_2. \tag{6.6}$$

Clearly $F(\delta_0) = 0$ and $F(\delta_i) = iL/(L-1) > 0$ for $i = 1, \dots, C$.

We use a basis of $\mathbb{R}^{C+1}$ such that $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_C)$ in the canonical basis is represented by $(\beta_0, \beta_1, \dots, \beta_C)$. We take $\beta_0 = \langle\alpha\rangle = \sum_{k=0}^C k\alpha_k$. For $C = 1$, $\langle\alpha\rangle_2 = \langle\alpha\rangle$ and $F(\alpha) = \beta_0(\beta_0 + 1/(L-1))$ vanishes uniquely for $\alpha = \delta_0$ in $\mathscr{P}$. If $C > 1$, we further take $\beta_1 = \langle\alpha\rangle_2 - (1 + 1/(L-1))\langle\alpha\rangle$, and then $F(\alpha) = (\beta_0)^2 - \beta_1$, hence the equation $F(\alpha) = 0$ defines in $\mathbb{R}^{C+1}$ a cylinder with parabolic base $(\beta_0)^2 = \beta_1$ in the appropriate two-dimensional vector subspace, $(\beta_2, \dots, \beta_C)$ being free to range over a $C-1$ dimensional subspace. This cylinder delimitates a convex open set $\{F(\alpha) < 0\}$, which has a (possibly empty) convex open intersection with $\mathscr{P}^o$.

A vector space X admits a decomposition $Y + Z$ if Y and Z are vector subspaces such that for any x in X, there are unique y in Y and z in Z with $x = y + z$. The decomposition is orthogonal for a quadratic form Q if we have $y^* Q z = 0$ for any column vectors y in Y and z in Z. We denote by $\operatorname{vect}(x)$ the line $\mathbb{R}x = \{ax\colon a \in \mathbb{R}\}$ generated by the vector x.

THEOREM 6.3. *For any $\alpha \in \mathscr{P}^o$ the decomposition $\mathscr{T} = \mathscr{N} + \mathrm{vect}(\bar{\alpha} - \alpha)$ is orthogonal for $Q(\alpha)$. The restriction of $Q(\alpha)$ to $\mathscr{N}$ coincides with the restriction of* $\mathrm{diag}(\alpha)^{-1}$ *and hence is positive definite* (*of rank* $C-1$), *and*

$$(\bar{\alpha} - \alpha)^* Q(\alpha)(\bar{\alpha} - \alpha) = \left(1 - \frac{L-1}{L}\frac{\mathrm{Var}(\alpha)}{\langle \alpha \rangle}\right)\frac{\mathrm{Var}(\alpha)}{\langle \alpha \rangle^2} = F(\alpha)\frac{L-1}{L}\frac{1}{\langle \alpha \rangle}\frac{\mathrm{Var}(\alpha)}{\langle \alpha \rangle^2}.$$

PROOF. We represent $\mathscr{T}$ in the basis such that $h \in \mathscr{T}$ has coordinates $(h_1, \dots, h_C)$, with $h_0 = -(h_1 + \cdots + h_C)$. Since $\mathrm{diag}(\alpha)^{-1}$ is positive definite, the restriction of the corresponding quadratic form on $\mathscr{T}$ is positive definite. We denote by B^{-1} its matrix. Then

$$h^* \mathrm{diag}(\alpha)^{-1} h = \sum_{i=0}^{C} \frac{1}{\alpha\{i\}} h_i^2 = \frac{1}{\alpha\{0\}}(h_1 + \cdots + h_C)^2 + \sum_{i=1}^{C} \frac{1}{\alpha\{i\}} h_i^2, \qquad h \in \mathscr{T}$$

and setting $u = (1, \dots, 1)^*$, with C terms, and $A = \mathrm{diag}(\alpha\{1\}, \dots, \alpha\{C\})$, we have

$$B^{-1} = A^{-1} + \frac{1}{\alpha\{0\}} uu^*, \qquad B = A - Au(Au)^* = A - Auu^*A, \tag{6.7}$$

where $BB^{-1} = I_C$ is easily checked using $u^*Au = \alpha\{1\} + \cdots + \alpha\{C\} = 1 - \alpha\{0\}$.

We set $w = (1, \dots, C)^*$ in this basis of $\mathscr{T}$. Then (6.3) implies

$$Q(\alpha) = B^{-1} - \frac{L-1}{L}\frac{1}{\langle \alpha \rangle} ww^* \tag{6.8}$$

and the restrictions of Q and B^{-1} to $w^\perp$ are equal. Since $w^*Bw > 0$, then $Bw \notin w^\perp$ for the canonical scalar product in this basis. Since $x \in w^\perp$ implies $x^*Qx = x^*B^{-1}x$ and $x^*QBw = 0$, the decomposition of $\mathscr{T}$ into $w^\perp + \mathrm{vect}(Bw)$ is orthogonal for Q. Then

$$Bw = Aw - (u^*Aw)Au = Aw - \langle \alpha \rangle Au = \langle \alpha \rangle (\bar{\alpha} - \alpha),$$

$$(Bw)^* QBw = \left(1 - \frac{L-1}{L}\frac{1}{\langle \alpha \rangle} w^*Bw\right) w^*Bw,$$

$$w^*Bw = w^*Aw - (u^*Aw)^2 = \langle \alpha \rangle_2 - \langle \alpha \rangle^2 = \mathrm{Var}(\alpha).$$

The theorem follows by writing down explicitly these terms, in particular $w^\perp = \mathscr{N}$. □

We now finish describing the function J. A subset K is uniformly convex in a vector space E with norm $\|\cdot\|$ if there is $k > 0$ such that

$$\inf_{x \in E - K} \left\| x - \frac{a+b}{2} \right\| \geq k\|a - b\|, \qquad a, b \in K.$$

A function f from $\mathbb{R}^d$ to $\mathbb{R}$ is uniformly convex if its epigraph $\{(x, z) \in \mathbb{R}^d \times \mathbb{R}\colon f(x) \leq z\}$ is uniformly convex, and is locally uniformly convex at x if its restriction to an open ball containing x is uniformly convex. If f is twice differentiable at x and its second differential matrix at x has strictly positive eigenvalues, then f is locally uniformly convex at x.

THEOREM 6.4. *In the independent case $L = 1$, $J(\cdot) = H(\cdot|q_\rho)$ is uniformly convex in $\mathscr{P}$. When $C = 1$ and $L \geq 2$, or $C = 2$ and $L = 2$, Q is positive definite and J is uniformly convex in $\mathscr{P}$. When $C \geq 2$ and $L \geq 3$, or $C \geq 3$ and $L \geq 2$, Q has signature $(C-1, 1)$ in the nonempty convex open set $\{F(\alpha) < 0\} \cap \mathscr{P}^o$, has signature $(C-1, 0)$ on the parabolic cylinder $\{F(\alpha) = 0\} \cap \mathscr{P}^o$ and has signature $(C, 0)$ in the nonempty open set $\{F(\alpha) > 0\} \cap \mathscr{P}^o$, at each point of which J is locally uniformly convex.*

PROOF. If $C = 1$, then $\langle \alpha \rangle_2 = \langle \alpha \rangle$ and $F(\alpha) > 0$ in $\mathscr{P}^o$. For $C = 2$ we have

$$F(\alpha) = \alpha\{1\}^2 + 4\alpha\{1\}\alpha\{2\} + 4\alpha\{2\}^2 + \frac{1}{L-1}\alpha\{1\} - \left(2 - \frac{2}{L-1}\right)\alpha\{2\}.$$

Thus if $L = 2$, then $F(\alpha) > 0$ in $\mathscr{P}^o$ and if $L \geq 3$, then

$$F(\alpha) \leq \alpha\{1\}(\alpha\{1\} + 4\alpha\{2\} + 1/2) - \alpha\{2\}(1 - 4\alpha\{2\})$$

and $F(\alpha) < 0$ for $0 < \alpha\{2\} < 1/4$ and $\alpha\{1\}$ sufficiently small. For $C = 3$ and $L \geq 2$,

$$\begin{aligned} F(\alpha) &\leq \langle \alpha \rangle^2 + 2\langle \alpha \rangle - \langle \alpha \rangle_2 = (\alpha\{1\} + 2\alpha\{2\} + 3\alpha\{3\})^2 + \alpha\{1\} - 3\alpha\{3\} \\ &\leq (\alpha\{1\} + 2\alpha\{2\})^2 + 6(\alpha\{1\} + 2\alpha\{2\})\alpha\{3\} + \alpha\{1\} \\ &\qquad - 3\alpha\{3\}(1 - 3\alpha\{3\}) \end{aligned}$$

and $F(\alpha) < 0$ for $0 < \alpha\{3\} < 1/3$ and $\alpha\{1\}$ and $\alpha\{2\}$ sufficiently small. We conclude for all $C \geq 3$ and $L \geq 2$ by a continuity argument. $\square$

We now obtain a good understanding of the behavior of J near q_ρ.

THEOREM 6.5. *The quadratic form $Q(q_\rho)$ is definite positive* (*of rank C*) *on $\mathscr{T}$, and the rate function J is locally uniformly convex at q_ρ.*

PROOF. We obtain, using (2.2),

$$\begin{aligned} \langle q_\rho \rangle_2 &= \sum_{k=1}^{C} k^2 q_\rho\{k\} = \rho\big(1 - q_\rho\{C\}\big)^{L-1} \sum_{k=1}^{C} k q_\rho\{k-1\} \\ &= \rho\big(1 - q_\rho\{C\}\big)^{L-1}\Big(\langle q_\rho \rangle - C q_\rho\{C\} + 1 - q_\rho\{C\}\Big) \end{aligned}$$

and using (2.6) and (6.6),

$$\begin{aligned} F(q_\rho) &= \rho\big(1 - q_\rho\{C\}\big)^{L-1}\Bigg((1 - q_\rho\{C\})\langle q_\rho \rangle + \frac{1}{L-1}(1 - q_\rho\{C\}) \\ &\qquad\qquad - \langle q_\rho \rangle + C q_\rho\{C\}\Bigg) \\ &= \rho\big(1 - q_\rho\{C\}\big)^{L-1}\Bigg((C - \langle q_\rho \rangle) q_\rho\{C\} + \frac{1}{L-1}(1 - q_\rho\{C\})\Bigg) > 0, \end{aligned} \tag{6.9}$$

where we conclude using the capacity constraint $\langle q_\rho \rangle \leq C$. $\square$

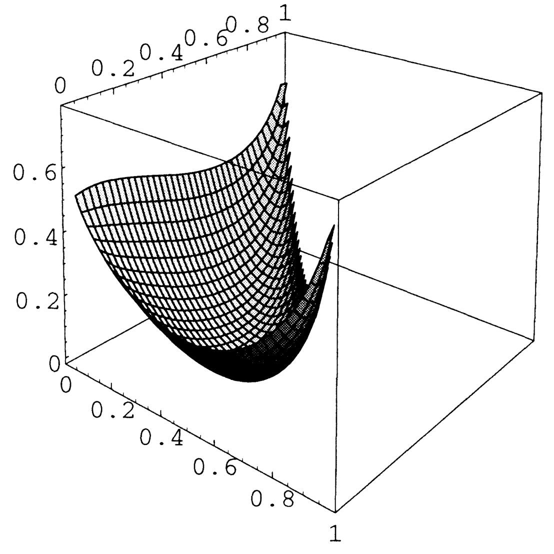

FIG. 1. $C = 2$, $\rho = 10$, $L = 5$.

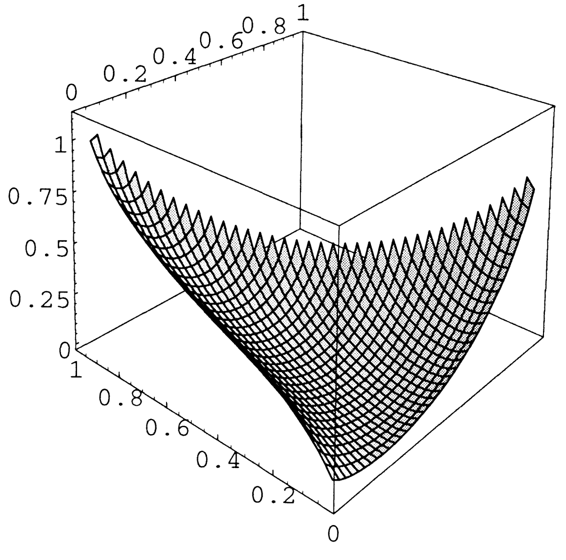

FIG. 2. $C = 2$, $\rho = 10$, $L = 50$.

THEOREM 6.6. *Let $\mathscr{T}$ be furnished with a norm $\|\cdot\|$ and $\mathscr{P}$ with the corresponding distance* (*the dimension being finite, all norms are equivalent*). *Let $B(q_\rho, \varepsilon)$ denote the ball of radius $\varepsilon > 0$ centered at q_ρ. Then*

$$\lim_{N\to\infty} \frac{1}{N} \log \pi^N(B(q_\rho, \varepsilon)^c) = -\theta\varepsilon^2 + O_{\varepsilon\to 0^+}(\varepsilon^3)$$

$$\text{for } \theta = \inf_{h\in\mathscr{T}:\ \|h\|=1} h^* Q(q_\rho)h > 0.$$

PROOF. Since J is nonnegative continuous on $\mathscr{P}$, vanishes only at q_ρ and is convex in a neighborhood of q_ρ, then for $\varepsilon > 0$ small enough the infimum of $J(\alpha)$ for $\alpha \notin B(q_\rho, \varepsilon)$ will be attained at the boundary. A Taylor expansion gives for any vector h of norm 1 $J(q_\rho + \varepsilon h) = \varepsilon^2 h^* Q(q_\rho)h + O_{\varepsilon\to 0^+}(\varepsilon^3)$. □

7. Some pictures. We present plots of the rate function J for $C = 2$ and $\rho = 10$, as a function of $\alpha\{1\}$ and $\alpha\{2\}$. In Figure 1, $L = 5$ and in Figure 2, $L = 50$. The nonconvexity is quite apparent.

In Figure 1 the origin is at the lower left, the $\alpha\{1\}$ axis points to the right and the $\alpha\{2\}$ axis points to the rear. In Figure 2 we have rotated the perspective, and the origin is the bottommost point (at the middle), the $\alpha\{1\}$ axis points to the right and the $\alpha\{2\}$ axis points to the left.

REFERENCES

[1] CHANG, C.-S. and WANG, H.-J. (1997). Large deviations for large capacity loss networks with fixed routing and polyhedral admission sets. *Discrete Event Dynamical Systems: Theory and Applications* **7** 391–418.

[2] DEMBO, A. and ZEITOUNI, O. (1993). *Large Deviation Techniques and Applications*. Jones and Bartlett, Boston.

[3] GRAHAM, C. and MÉLÉARD, S. (1994). Chaos hypothesis for a system interacting through shared resources. *Probab. Theory Related Fields* **100** 157–173.

[4] GRAHAM, C. and MÉLÉARD, S. (1995). Dynamic asymptotic results for a generalized star-shaped loss network. *Ann. Appl. Probab.* **5** 666–680.

[5] GRAHAM, C. and MÉLÉARD, S. (1997). An upper bound of large deviations for a generalized star-shaped loss network. *Markov Processes and Related Fields* **3** 199–223.

[6] GRAHAM, C. and MÉLÉARD, S. (1997). A large deviation principle for a large star-shaped loss network with links of capacity one. *Markov Processes and Related Fields* **3** 475–492.

[7] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3d ed. Wiley, New York.

[8] HUNT, P. J. (1995). Loss networks under diverse routing, the symmetric star network. *Adv. in Appl. Probab.* **25** 255–272.

[9] KELLY, F. P. (1991). Loss networks. *Ann. Appl. Probab.* **1** 319–378.

[10] O'Connell, N. (1997). From laws of large numbers of large deviation principles. *Markov Processes and Related Fields* **3** 475–492.

[11] O'CONNELL, N. (1998). Some large deviation results for sparse graphs. *Probab. Theory Related Fields* **110** 277–185.

[12] O'CONNELL, N. (1998). A large deviations heuristic made precise. Techinical Report HPL-BRIMS-9810.

[13] WHITT, W. (1985). Blocking when service is required from several facilities simultaneously. *AT&T Technical Journal* **64** 1807–1856.

[14] ZIEDINS, I. B. and KELLY, F. P. (1989). Limit theorems for loss networks with diverse routing. *Adv. in Appl. Probab.* **21** 804–830.

CMAP, ÉCOLE POLYTECHNIQUE
91128 PALAISEAU CEDEX
FRANCE
E-mail: Carl@cmapx.polytechnique.fr

BRIMS, HEWLETT-PACKARD LABS
FILTON ROAD, STOKE GIFFORD
BRISTOL BS34 8QZ
UNITED KINGDOM
E-mail: noc@hplb.hpl.hp.com