The Annals of Applied Probability
1996, Vol. 6, No. 2, 577–616

R-POSITIVITY, QUASI-STATIONARY DISTRIBUTIONS AND RATIO LIMIT THEOREMS FOR A CLASS OF PROBABILISTIC AUTOMATA

BY P. A. FERRARI, H. KESTEN AND S. MARTÍNEZ

Universidade de São Paulo, Cornell University and Universidad de Chile

We prove that certain (discrete time) probabilistic automata which can be absorbed in a "null state" have a normalized quasi-stationary distribution (when restricted to the states other than the null state). We also show that the conditional distribution of these systems, given that they are not absorbed before time n, converges to an honest probability distribution; this limit distribution is concentrated on the configurations with only finitely many "active or occupied" sites.

A simple example to which our results apply is the discrete time version of the subcritical contact process on $\mathbb{Z}^d$ or oriented percolation on $\mathbb{Z}^d$ (for any $d \geq 1$) as seen from the "leftmost particle." For this and some related models we prove in addition a central limit theorem for $n^{-1/2}$ times the position of the leftmost particle (conditioned on survival until time n).

The basic tool is to prove that our systems are R-positive-recurrent.

1. Introduction and principal results. Let $\{X_n\}_{n\geq 0}$ be a Markov chain on a countable state space S, with an absorbing state s_0. We shall deal exclusively with the discrete time case in this paper, but we believe that all the results and proofs have analogues in the continuous time case. We shall write S_0 for $S \setminus \{s_0\}$. As usual, we denote the probability measure governing $X_\bullet$ when conditioned to start at $X_0 = x$ by P_x. Then for any probability measure ν on S,

$$P_\nu = \sum_{x\in S} \nu(x) P_x = \nu P$$

is the measure which governs $X_\bullet$ when the initial distribution is ν. Let E_x and E_ν denote expectation with respect to P_x and P_ν, respectively. The transition probabilities are

$$P(x, y) = P_x\{X_1 = y\} \quad \text{and} \quad P^n(x, y) = P_x\{X_n = y\}. \tag{1.1}$$

It is convenient to introduce the restriction $\widehat{P}$ of P to $S_0 \times S_0$. Because s_0 is absorbing, we then also have

$$(\widehat{P})^n(x, y) = P_x\{X_n = y\} = P^n(x, y) \quad \text{for } x, y \in S_0.$$

Received October 1994; revised May 1995.

AMS 1991 *subject classifications*. Primary 60J10, 60F05; secondary 60K35.

Key words and phrases. Absorbing Markov chain, quasi-stationary distribution, ratio limit theorem, Yaglom limit, R-positivity, central limit theorem.

Throughout we assume that

(1.2) $\widehat{P}$ is irreducible; that is, for all $x, y \in S_0$ there exists an $n = n(x, y)$ with $P^n(x, y) > 0$.

The *absorption time* is

$$T = T(s_0) = \inf\{n \geq 0\colon\ X_n = s_0\} \tag{1.3}$$

and we assume that absorption is certain. This means that for some $x \in S_0$ (and hence all x)

$$P_x\{T < \infty\} = 1. \tag{1.4}$$

The fact that s_0 is absorbing means of course that

$$X_n = s_0 \quad \text{for all } n \geq T \quad \text{and} \quad P^n(s_0, x) = \delta(s_0, x), \qquad n \geq 0. \tag{1.5}$$

A *normalized quasi-stationary distribution* for X is a *probability* measure ν on S_0 which satisfies the invariance condition

$$\nu \widehat{P}^n = r(n)\nu, \tag{1.6}$$

where, necessarily,

$$r(n) = \sum_{y \in S_0} \nu \widehat{P}^n(y) = P_\nu\{T > n\}. \tag{1.7}$$

Note that these distributions are conditionally invariant, in the sense that

$$P_\nu\{X_n = y | T > n\} = \nu(y) \quad \text{for all } y \in S_0.$$

A normalized quasi-stationary distribution ν_0 is called *minimal* if

$$E_{\nu_0} T = \inf\{E_\nu T\colon\ \nu \text{ a normalized quasi-stationary distribution}\}. \tag{1.8}$$

The interest in normalized quasi-stationary distributions arises from the fact that if for some initial distribution ν on S_0,

$$\frac{P_\nu\{X_n = y\}}{P_\nu\{T > n\}} \to \mu(y), \qquad y \in S_0, \tag{1.9}$$

for some probability distribution μ on S_0, then μ is necessarily a normalized quasi-stationary distribution. [See Seneta and Vere-Jones (1966), Theorem 4.1]. If the limit in (1.9) exists, we shall call it a Yaglom limit, because Yaglom (1947) proved the existence of this limit for subcritical branching processes (when ν is concentrated on one point x). Such a limit μ is also called a *conditional limit distribution* in the literature. There is an extensive literature discussing the existence of normalized quasi-stationary distributions and the Yaglom limit; see Ferrari, Kesten, Martínez and Picco (1995) for some references. An additional recent reference is Roberts and Jacka (1994). Recently, Ferrari, Kesten, Martínez and Picco (1995) gave a necessary and sufficient condition for the existence of a normalized quasi-stationary distribution for chains $X_\bullet$ which satisfy, in addition to (1.2) and (1.4), the condition

$$\lim_{x\to\infty} P_x\{T \le t\} = 0 \quad \text{for all fixed } t < \infty. \tag{1.10}$$

This of course means that for all $t < \infty$ and $\varepsilon > 0$, $P_x\{T \le t\} \le \varepsilon$ for all but finitely many x. [Actually Ferrari, Kesten, Martínez and Picco (1995) deal with the continuous time case, but their results carry over to discrete time; see Kesten (1995), Theorem A, for a statement of the discrete time result.] Unfortunately, (1.10) is rarely fulfilled for chains $X_\bullet$ which describe interacting particles which also have a spatial position or which can be of infinitely many types. The absorbing state is the state in which no particle is present. Typically in such models, the probability of absorption in unit time from any of the infinitely many states in which only one particle is present is bounded away from 0–infinitely many states, because the single particle can have infinitely many positions or types; the above phenomenon occurs if the probability for a particle to die in one time unit is uniformly bounded away from 0. A simple special example (which was, in fact, the principal motivation for this paper) is the subcritical contact process or oriented percolation on $\mathbb{Z}^d$ as seen from the "leftmost" particle (with state space S a certain collection of finite subsets of $\mathbb{Z}^d$ containing $\{0\}$, plus the empty set). Our aim here is to prove the existence of a normalized quasi-stationary distribution and a Yaglom limit for a class of interacting particle systems and probabilistic automata which includes oriented percolation on $\mathbb{Z}^d$, $d \ge 1$. This will be done by proving those chains R-positive-recurrent. (A more detailed description of some of these examples and the role of the leftmost particle is given before the statement of Theorem 2 below; full details are in Section 4.) Pakes (1995) investigated some other examples in which (1.10) may fail.

We remind the reader of some basic facts which hold solely under the assumption (1.2) [see Vere-Jones (1967)]:

(1.11) The period $p := \text{g.c.d.}\ \{n\colon\ \widehat{P}^n(x, x) > 0\}$ is finite and is the same for all $x \in S_0$;

(1.12) S_0 can be decomposed into p disjoint subclasses $S_{0,0}, \dots, S_{0,p-1}$ so that $\widehat{P}(x, y) > 0$ only if $x \in S_{0,j}$, $y \in S_{0,j+1}$ for some $0 \le j \le p-1$ (here and in the sequel we take $S_{0,j_1} = S_{0,j_2}$ when $j_1 \equiv j_2 (\text{mod}\ p)$; the $S_{0,j}$ are referred to as the *cyclically moving subclasses* or *periodic subclasses*);

(1.13) if $x \in S_{0,i}$ and $y \in S_{0,j}$, then $1/R := \lim_{n\to\infty}[\widehat{P}^{np+j-i}(x, y)]^{1/np}$ exists and the value of R is independent of x, y, i and j;

$$1 \le R < \infty. \tag{1.14}$$

The chain $X_{\bullet}$ and $\widehat{P}$ are called *R-recurrent* if for some $x \in S_0$,

$$\sum_{n=0}^{\infty} R^{np} \widehat{P}^{np}(x, x) = \infty, \tag{1.15}$$

and it is called *R-positive-recurrent* if, in addition, for some $x \in S_0$,

$$\limsup_{n \to \infty} R^{np} \widehat{P}^{np}(x, x) > 0. \tag{1.16}$$

Again (1.15) and (1.16) either hold for all x in S_0 simultaneously or for no x. If (1.15) and (1.16) hold, then there exist functions f: $S_0 \to (0, \infty)$, μ: $S_0 \to (0, \infty)$ which satisfy

$$\widehat{P}^n f(x) = \sum_{y \in S_0} \widehat{P}^n(x, y) f(y) = R^{-n} f(x), \qquad x \in S_0,\ n \geq 0, \tag{1.17}$$

$$\mu \widehat{P}^n(y) = \sum_{x \in S_0} \mu(x) \widehat{P}^n(x, y) = R^{-n} \mu(y), \qquad y \in S_0,\ n \geq 0, \tag{1.18}$$

$$\sum_{x \in S_{0,j}} f(x)\mu(x) = 1, \qquad 0 \leq j \leq p-1, \qquad \sum_{x \in S_0} f(x)\mu(x) = p \tag{1.19}$$

and for $x \in S_{0,i}$, $y \in S_{0,j}$ for some $0 \leq i, j \leq p-1$,

$$\lim_{n \to \infty} R^{np+j-i} \widehat{P}^{np+j-i}(x, y) = f(x)\mu(y). \tag{1.20}$$

The functions f and μ are uniquely determined up to multiplicative constants by (1.17) and (1.18). Equation (1.18) does not say that μ is a (multiple of a) normalized quasi-stationary distribution, because $\sum_{x \in S_0} \mu(x)$ may diverge. In order to obtain a normalized quasi-stationary distribution and for the Yaglom limit relation (1.9) we need

$$\sum_{x \in S_0} \mu(x) < \infty \tag{1.21}$$

[see Seneta and Vere-Jones (1966), Theorem 3.1]. It is an important step in our proof to show that (1.21) indeed holds under the conditions of Theorem 1 below, which is our principal result.

THEOREM 1. *Assume that the Markov chain* $\{X_n\}_{n \geq 0}$ *satisfies* (1.2) *and* (1.4) *as well as the following conditions*:

(1.22) *there exist a nonempty set* $\mathscr{U}_1 \subset S_0$, *an* $\varepsilon_0 > 0$ *and a constant* C_1 *such that for all* $x \in \mathscr{U}_1$ *and all* $n \geq 0$, $P_x\{T > n$, *but* $X_l \notin \mathscr{U}_1$ *for all* $1 \leq l \leq n\} \leq C_1(R + \varepsilon_0)^{-n}$,

(1.23) *there exist a state* $x_0 \in \mathscr{U}_1$ *and a constant* C_2 *such that for all* $x \in \mathscr{U}_1$ *and* $n \geq 0$, $P_x\{T > n\} \leq C_2 P_{x_0}\{T > n\}$

and

(1.24) *there exist a finite set* $\mathscr{U}_2 \subset S_0$ *and constants* $0 \le n_0 < \infty$, $C_3 > 0$, *such that for all* $x \in \mathscr{U}_1$, $P_x\{X_n \in \mathscr{U}_2$ *for some* $n \le n_0\} \ge C_3$.

Then $X_\bullet$ *is R-positive-recurrent, so that* (1.17)–(1.20) *hold for* $p =$ *period of* $\widehat{P}$ *and some strictly positive functions* f *and* μ. *Moreover*, (1.21) *is satisfied and*

$$\widetilde{\mu}(x) := \frac{\mu(x)}{\sum_{y\in S_0}\mu(y)}, \qquad x \in S_0, \tag{1.25}$$

defines a minimal quasi-stationary distribution for $X_\bullet$. *Also, if* $S_{0,0}, \dots, S_{0,p-1}$ *are the cyclically moving subclasses as in* (1.12), *and* $x \in S_{0,i}$, *then*

$$\lim_{\substack{n\to\infty\\ n\equiv j\,(\mathrm{mod}\,p)}} R^n P_x\{T > n\} = f(x)\sum_{y\in S_{0,i+j}}\mu(y) \tag{1.26}$$

and for $x \in S_{0,i}$, $y \in S_{0,j}$,

$$\lim_{\substack{n\to\infty\\ n\equiv j-i\ (\mathrm{mod}\,p)}} P_x\{X_n = y | T > n\} = \frac{\mu(y)}{\sum_{v\in S_{0,j}}\mu(v)}. \tag{1.27}$$

Finally, for each $x \in S_0$ *there exist an* $\eta(x) > 0$ *and a* $C_4(x) < \infty$ *such that*

$$P_x\{X_{np} = x,\ \text{but } X_s \ne x \text{ for } 1 \le s \le np-1\} \le C_4(x)\frac{1}{(R+\eta(x))^n}. \tag{1.28}$$

REMARK 1. Equation (1.27) can be easily generalized to give convergence of finite-dimensional distributions. That is, for $0 < s_1 < \cdots < s_l$, $x \in S_{0,i}$, $y_1 \in S_{0,j_1}, \dots, y_l \in S_{0,j_l}$ it holds that

$$\begin{aligned}\lim_{n\to\infty} P_x\big\{X_{\lfloor s_r n\rfloor p + j_r - i} = y_r,\ 1 \le r \le l \big| T > \lfloor s_l n\rfloor p + j_l - i\big\}\\ = \frac{\mu(y_l)}{\sum_{v\in S_{0,j_l}}\mu(v)}\prod_{r=1}^{l-1}[f(y_r)\mu(y_r)].\end{aligned} \tag{1.29}$$

REMARK 2. A consequence of (1.29) and (1.19) is that for all $\varepsilon > 0$ and $0 < s \le 1$, there exists a *finite* set $S(\varepsilon, s) \subset S$ such that for all $x \in S_0$,

$$\lim_{n\to\infty} P_x\big\{X_{\lfloor sn\rfloor} \in S(\varepsilon, s) | T > n\big\} \ge 1 - \varepsilon.$$

In other words, when conditioned on no absorption until time n, then the chain $X_\bullet$ tends to spend "most of its time near s_0." This contrasts with some of the examples in Ferrari, Kesten, Martínez and Picco (1995). For example, let $\widetilde{X}_n$ be the absorbing random walk on $\mathbb{Z}_+$ with 0 as absorbing state and which moves from $x \ge 1$ to $x+1$ or $x-1$ with probability p and $q = 1-p$, respectively, with $0 < p < \frac{1}{2}$. Then $\{n^{-1/2}\widetilde{X}_{\lfloor sn\rfloor}\}_{0\le s\le 1}$ conditioned on $T > n$

essentially behaves like a positive Brownian excursion (scaled so that it has length 1). In particular, for fixed $0 < s < 1$,

$$P_x\{|\widetilde{X}_{\lfloor sn\rfloor} < \varepsilon\sqrt{n}|T > n\}$$

can be made small uniformly in n, by taking ε small. Thus $\widetilde{X}$ conditioned on $T > n$ spends most of its time far away from the absorbing state 0.

Theorem 1 has some value only if its hypotheses can be verified without too much trouble for a reasonable class of examples. The hypothesis which looks the most troubling is (1.22). The next theorem shows that this actually follows directly from known results in many instances of so-called probabilistic cellular automata, which are similar to the subcritical contact process. We begin with a Markov chain $\{\widetilde{X}_n\}_{n\geq 0}$ on a countable state space

(1.30) $\widetilde{S} = \widetilde{S}_{d,\kappa}$ which is a subset of the collection of all functions $x\colon \mathbb{Z}^d \to \{0, 1, \dots, \kappa - 1\}$ with $x(z) \neq 0$ for only finitely many z.

Here $\kappa \geq 2$ is some fixed integer. Note that $\mathbb{Z}^d$ can be replaced by other lattices and this should indeed be done to treat the usual oriented percolation. For simplicity we restrict ourselves here to the state space (1.30). Let the state $\underline{0}$ with all components equal to 0 be the absorbing state. Thus

$$P\{\widetilde{X}_n = x|\widetilde{X}_0 = \underline{0}\} = \delta(\underline{0}, x), \qquad n \geq 0,\ x \in \widetilde{S}. \tag{1.31}$$

The space $\widetilde{S}$ and the transition probabilities have to be such that

(1.32) the transition probability matrix is irreducible on $\widetilde{S}_0 = \widetilde{S} \setminus \{\underline{0}\}$.

(It is for this reason that $\widetilde{S}$ may have to be taken only as a *subset* of

$$\{x \in \{0, 1, \dots, \kappa - 1\}^{\mathbb{Z}^d}\colon\ x(z) \neq 0 \text{ for only finitely many } z\},$$

rather than the full set.) We further impose the following fairly common conditions on the transition probabilities:

(1.33) *Translation invariance*: $P\{\widetilde{X}_1 = y \oplus u|\widetilde{X}_0 = x \oplus u\} = P\{\widetilde{X}_1 = y|\widetilde{X}_0 = x\}$ for all $u \in \mathbb{Z}^d$, where the state $x \oplus u$ is specified by $(x \oplus u)(z) = x(z + u)$, $z \in \mathbb{Z}^d$, and similarly for $y \oplus u$;

(1.34) *Independence of coordinates*: Conditionally on $\widetilde{X}_0, \dots, \widetilde{X}_n$, the coordinates $\{\widetilde{X}_{n+1}(z)\colon z \in \mathbb{Z}^d\}$ of $\widetilde{X}_{n+1}$ are independent;

(1.35) *Finite range*: There exists some $\rho < \infty$ such that the conditional distribution of $\widetilde{X}_{n+1}(z)$, given $\widetilde{X}_n$, depends only on $\widetilde{X}_n(z')$ for $\|z' - z\|_2 \leq \rho$.

It is trivial to check that the discrete time version of the contact process and oriented percolation (see Section 4) satisfy (1.31)–(1.35) for $\kappa = 2$. In analogy with the terminology for that chain we shall call sites z with $x(z) = 0$ ($\neq 0$) *vacant* (and *occupied*, respectively) in state x. It is convenient to think of a particle sitting at each occupied site, the particle having one of the types $1, \dots, \kappa - 1$. The chain $\widetilde{X}_\bullet$ then describes motion, birth, death and change of type of a finite system of particles.

Theorem 1 will not apply to $\widetilde{X}$ itself. As illustrated by oriented percolation, we cannot expect the limit in (1.27) to have a nonzero value because the occupied sites of $\widetilde{X}_n$ will wander all over the space $\mathbb{Z}^d$ (even under the condition $\{T > n\}$). Also condition (1.24) will fail because particles cannot jump in fixed time n_0 from far away to a site which is occupied in one of the finitely many states in $\mathscr{U}_2$. Theorem 1 will only apply to the *relative* positions of the occupied sites. To make this precise, let $\xi(n) = \{\xi_1(n), \dots, \xi_\nu(n)\}$ be the occupied sites at time n, that is the sites in $\mathbb{Z}^d$ with $\widetilde{X}_n(\xi_i(n)) \neq 0$. Here $\nu = \nu(n)$ is random. If no sites are occupied at time n, we take $\nu(n) = 0$, $\xi(n) = \phi$. If $\nu > 0$, the order of the $\xi_i(n)$ is arbitrary, except that we take ξ_1 to be the first among $\xi_1, \dots, \xi_\nu$ in the lexicographical ordering on $\mathbb{Z}^d$. We call $\xi_1(n)$ the *leftmost occupied position* at time n. Finally define

$$X_n = \begin{cases} \underline{0}, & \text{if } \nu(n) = 0, \\ \widetilde{X}_n \oplus (-\xi_1(n)). \end{cases} \tag{1.36}$$

Thus $X_n(z) = j$ precisely when $\widetilde{X}_n(z + \xi_1(n)) = j$. From the translation invariance (1.33) it is not hard to see that $\{X_n\}_{n \geq 0}$ is again a Markov chain (this is the only place where translation invariance is strictly needed). The state space S of the X-chain is of the same form as that of $\widetilde{X}$. Only $X_n(0) \neq 0$ whenever $X_n \neq \underline{0}$, and $X_n(z) \neq 0$ can only occur at a z, which comes after 0 in the lexicographical ordering of $\mathbb{Z}^d$. The absorbing state for $X_\bullet$ is again $\underline{0}$.

Here are some final definitions. For $x \in \widetilde{S}$ or $x \in S$,

$$\begin{aligned} |x| &= \text{number of occupied sites in } x \\ &= \text{number of } z \in \mathbb{Z}^d \text{ with } x(z) \neq 0. \end{aligned} \tag{1.37}$$

The use of the same symbol $|\ |$ for points in $\widetilde{S}$ and in S will not lead to any serious confusion, since in both cases $|x|$ represents the number of occupied sites. For similar reasons we shall use T for the absorption time of $\widetilde{X}$ and of X. The process $\{\widetilde{X}_n\}$ is called *strongly subcritical* if

$$n^d \sup_{\substack{x \in \widetilde{S} \\ z \in \mathbb{Z}^d}} P_x\{\widetilde{X}_n(z) \neq 0\} \to 0 \quad \text{as } n \to \infty. \tag{1.38}$$

THEOREM 2. *Let $\widetilde{X}$ be a Markov chain on the space $\widetilde{S}$ of* (1.30) *whose transition probabilities satisfy* (1.31)–(1.35). *Let X_n be defined by* (1.36). *Assume* *that $\{\widetilde{X}_n\}$ is strongly subcritical. Then for each $\gamma > 0$ there exists a $K < \infty$ and a $C_5 = C_5(\gamma) < \infty$ such that for all x with $|x| \le K$,*

$$P_x\{T > n,\ \text{but } |X_l| > K \text{ for all } 1 \le l \le n\} \le C_5\gamma^n, \qquad n \ge 0. \tag{1.39}$$

Clearly (1.39) gives us (1.22) with $\mathscr{U}_1 = \{x: |x| \le K\}$ by taking $\gamma < R^{-1}$, say $\gamma = (R+1)^{-1}$. Thus, Theorem 2 reduces checking (1.22) to checking that our process is strongly subcritical. For some processes it is known that if they die out w.p.1 [i.e., satisfy (1.4)], then they are automatically strongly subcritical. In other examples it should be possible to use a Peierls argument or simple comparisons with other chains to show that the chain $X_\bullet$ is strongly subcritical, at least for certain choices of the transition probabilities. (See also the discussion at the end of Example A in Section 4.) Condition (1.24) is usually innocuous; as illustrated in the proof of Proposition 1, it just depends on probabilities of particles disappearing being bounded away from 0. Condition (1.23), which is trivial for the contact process or oriented percolation (see Remark 8 in Section 4), turns out to be nontrivial in general. We verify it in Section 4 for some chains of the form described above. For the special case of the discrete time contact process or oriented percolation, all hypotheses are verified in this same section.

Finally, in Section 5 we derive a central limit theorem for the absolute location of the occupied sites in the $\widetilde{X}$-chain. The conclusion of Theorem 1 tells us that X_n, conditioned on $\{T > n\}$, has a limit distribution. Thus, for X_n as in (1.36), this tells us that the number of particles and their distances to the leftmost occupied position $\xi_1(n)$ are tight (under the condition $\{T > n\}$). However, the absolute position of the occupied set, by which we simply mean $\xi_1(n)$ itself, will not be tight without normalization. In fact our last theorem shows that $\xi_1(n)/\sqrt{n}$ is usually asymptotically normally distributed and independent of X_n (still under the condition $\{T > n\}$). This result follows almost immediately from standard central limit theorems for recurrent Markov chains, by means of the strong recurrence result (1.28).

THEOREM 3. *Let $\{X_n\}$ be defined by* (1.36) *for a chain $\widetilde{X}_\bullet$ with rates satisfying* (1.31)–(1.35). *Assume that* (1.22)–(1.24) *are satisfied. Then there exist an $M \in \mathbb{R}^d$ and a $d \times d$ matrix Σ such that*

$$\frac{\xi_1(n) - nM}{\sqrt{n}} \text{ conditioned on } \{T > n\} \text{ converges in distribution to } N(\mathbf{0}, \Sigma). \tag{1.40}$$

More generally let p be the period of (*the restriction to S_0 of*) *$X_\bullet$ and let $S_{0,0}, \dots, S_{0,p-1}$ be the cyclically moving subclasses in S_0* [*cf.* (1.12)]. *Then, for $0 < s_1 < s_2 < \cdots < s_l$, $x \in S_{0,i}$, $y_1 \in S_{0,j_1}, \dots, y_l \in S_{0,j_l}$ and $\gamma_1, \dots, \gamma_l \in$*

$\mathbb{R}$, $w \in \mathbb{Z}^d$, *as* $n \to \infty$,

$$P\left\{ \frac{\xi_1(\lfloor s_r n \rfloor p + j_r - i) - s_r npM}{\sqrt{np}} \le \gamma_r,\ X_{\lfloor s_r n \rfloor p + j_r - i} = y_r,\ 1 \le r \le l \,\middle|\, X_0 = x,\ \xi_1(0) = w,\ T > \lfloor s_l n \rfloor p + j_l - i \right\} \tag{1.41}$$

$$\to P\{G_{s_r} \le \gamma_r,\ 1 \le r \le l\} \frac{\mu(y_l)}{\sum_{v \in S_{0,\,j_l}} \mu(v)} \prod_{r=1}^{l-1} [f(y_r)\mu(y_r)],$$

where $\{G_t\}_{t \ge 0}$ *is a d-dimensional Gaussian process with mean zero,* $G_0 = \mathbf{0}$ *and covariance function*

$$E\{G_s G_t^*\} = (s \wedge t)\Sigma$$

(G^* *is the transpose of the column vector* G). *In particular this result is valid for the subcritical discrete time contact process* (*site or bond version*) *on* $\mathbb{Z}^d$ *with* $M = \mathbf{0}$, $\Sigma = \sigma^2 \times (identitymatrixofdimensiond)$ *for some* $\sigma^2 > 0$.

If Σ is the zero matrix, then (1.41) means that for all $\varepsilon > 0$,

$$P\left\{ \frac{\|\xi_1(\lfloor s_r n \rfloor p + j_r - i) - s_r npM\|}{\sqrt{n}} \le \varepsilon,\ X_{\lfloor s_r n \rfloor p + j_r - i} = y_r,\ 1 \le r \le l \,\middle|\, X_0 = x,\ \xi_1(0) = w,\ T > \lfloor s_l n \rfloor p + j_l - i \right\}$$

$$\to \frac{\mu(y_l)}{\sum_{v \in S_{0,\,j_l}} \mu(v)} \prod_{r=1}^{l-1} [f(y_r)\mu(y_r)].$$

REMARK 3. Note that the condition $X_0 = x$, $\xi_1(0) = w$ fixes the full state of $\widetilde{X}_0$. If $X_0 = x$ and $\xi_1(0) = w$, then $\widetilde{X}_0(z) = x(z - w)$.

REMARK 4. Some more information on M and Σ is given in (5.12), (5.13) and succeeding lines.

REMARK 5. A central limit theorem for oriented percolation (in a Fourier form) was proven before by Nguyen and Yang (1995). They used lace expansions, which only work in high dimension and seem to be restricted to oriented percolation. It would be difficult to extend their method to the more general class of processes described before Theorem 2. On the other hand, Nguyen and Yang also treat the *critical* oriented percolation in addition to the subcritical one. Our method does not apply at all to the critical case.

For the more difficult case of ordinary, unoriented percolation, a related—but slightly different—central limit result was first proved in Campanino, Chayes and Chayes (1991).

REMARK 6. It is not difficult to extend the limit theorem (1.41) for the finite-dimensional distributions of $\{\xi_1(\lfloor sn\rfloor)/\sqrt{n}\}_{s\geq 0}$ conditioned on $\{T > n\}$ to a full invariance principle (of the ξ_1 coordinate only, though).

2. *R*-positive-recurrence. In this section we shall prove Theorem 1 for general absorbing Markov chains. Equations (1.2), (1.4) and (1.22)–(1.24) will be in force throughout this section.

Here is a brief outline of the proof. Fix a state y_0. Condition (1.22) [in conjunction with (1.13)] says more or less that, conditionally on $T > n$, the chain cannot stay away from $\mathscr{U}_1$ too long. This will be used to show that in fact $X_{\bullet}$ will visit $\mathscr{U}_1$ at a time near n, with a probability bounded away from 0. By (1.24), $X_{\bullet}$ will then also visit $\mathscr{U}_2$ and even y_0 at a time near n, with a probability bounded away from 0. We will actually show that even the probability for $X_k = y_0$ and $X_n = y_0$ is on the average (over $k \in [0, n]$) not too small, still under the condition $\{T > n\}$. A little renewal theory then shows that $X_{\bullet}$ must be R-positive-recurrent (see Lemma 5). As already stated, Vere-Jones' results then show that there exist f and μ satisfying (1.17)–(1.20). One can therefore introduce the honest transition probability matrix

$$P^*(x, y) = \frac{RP(x, y)f(y)}{f(x)}, \qquad x, y \in S_0.$$

This is positive-recurrent. From the fact that $\{X_n = y_0\}$ has not too small a probability, we will obtain that $P_{y_0}\{T > n\}$ is of the same order as $P^n(y_0, y_0)$. Using P^* and a Markov chain X^* with this transition probability quickly leads to (1.21), (1.26) and (1.27).

To prove (1.28) we appeal to the hypothesis (1.23). Specifically, this is used to show that the function f is bounded on $\mathscr{U}_1$. This will allow us to reduce (1.28) to fairly standard exponential bounds or large deviation estimates for the chain X^* (these no longer involve R).

The above steps are separated into various lemmas. The R-positive-recurrence and (1.17)–(1.20) are proven in Lemma 5; the fact that the $\tilde{\mu}$ of (1.25) is a minimal quasi-stationary distribution is in Lemma 7. Finally, (1.21) and (1.26)–(1.28) are proven in Lemmas 6 and 8. *We carry out the proof only for the aperiodic case* ($p = 1$); the extension to general p is routine.

Throughout D_i will stand for a strictly positive, finite constant whose value is of no significance to us; the same symbol D_i may represent a different constant in different proofs, but we will not vary the value of D_i in a single lemma. Also $C_1, C_2, \dots$ will be constants in $(0, \infty)$, but they maintain the same value throughout this section.

We begin with a simple reduction, based merely on the irreducibility assumption (1.2), which is not strictly needed, but which simplifies the formulae.

LEMMA 1. *For any fixed* $y_0 \in S_0$ *we may assume that* $\mathscr{U}_2$ *consists of* y_0 *only and that the* x_0 *of* (1.23) *equals* y_0. (*In particular we may assume that* $\mathscr{U}_1$ *contains* y_0.)

PROOF. Fix $y_0 \in S_0$. By the irreducibility of $\widehat{P}$ and the finiteness of $\mathscr{U}_2$ there exists an $m_0 < \infty$ and a $D_1 > 0$ such that for all $y \in \mathscr{U}_2$,

$$P^k(y, y_0) \geq D_1 \quad \text{for some } 0 \leq k \leq m_0. \tag{2.1}$$

Then, by (1.24), for $x \in \mathscr{U}_1$

$$\begin{aligned} P_x\{X_n = y_0 \text{ for some } n \leq n_0 + m_0\} \\ &\geq \sum_{j=0}^{n_0} \sum_{y \in \mathscr{U}_2} P_x\{X_{\bullet} \text{ first visits } \mathscr{U}_2 \text{ at time } j,\ X_j = y\} \\ &\qquad\qquad \times P_y\{X_k = y_0 \text{ for some } k \leq m_0\} \\ &\geq D_1 \sum_{j=0}^{n_0} P_x\{X_{\bullet} \text{ first visits } \mathscr{U}_2 \text{ at time } j\} \geq D_1 C_3. \end{aligned}$$

Thus, by replacing C_3 by $D_1 C_3$ and n_0 by $n_0 + m_0$, we may assume (1.24) valid with $\mathscr{U}_2 = \{y_0\}$.

Next we compare $P_{x_1}\{T > n\}$ and $P_{y_1}\{T > n\}$ for arbitrary $x_1, y_1 \in S_0$. These two probabilities will be of the same order. Indeed, there exists a $k = k(x_1, y_1)$ so that

$$P^k(y_1, x_1) > 0.$$

Since $X_k = x_1 \in S_0$ implies $T > k$, we also have

$$\begin{aligned} P_{y_1}\{T > n\} &\geq P^k(y_1, x_1) P_{x_1}\{T > n - k\} \\ &\geq P^k(y_1, x_1) P_{x_1}\{T > n\}. \end{aligned} \tag{2.2}$$

In particular, if y_0 is any fixed state, then the inequality in (1.23) will continue to hold (with a suitable change in C_2) if we replace x_0 by y_0. However, (1.23) in its original form required that $x_0 \in \mathscr{U}_1$, and y_0 may not belong to $\mathscr{U}_1$. In the latter case we simply add y_0 to $\mathscr{U}_1$ and verify that (1.22) and (1.24) remain valid after this enlargement of $\mathscr{U}_1$. To verify (1.22) we note that the irreducibility of $\widehat{P}$ shows that if y_0 does not belong to $\mathscr{U}_1$, then there exists some $k > 0$ and some $y \in \mathscr{U}_1$, such that

$$P_y\{X_k = y_0,\ X_l \notin \mathscr{U}_1 \text{ for } 1 \leq l \leq k\} > 0.$$

Moreover,

$$\begin{aligned} P_y\{T > n + k,\ \text{but } X_l \notin \mathscr{U}_1 \text{ for all } 1 \leq l \leq n + k\} \\ &\geq P_y\{X_k = y_0,\ X_l \notin \mathscr{U}_1 \text{ for } 1 \leq l \leq k\} \\ &\qquad \times P_{y_0}\{T > n,\ \text{but } X_l \notin \mathscr{U}_1 \text{ for all } 1 \leq l \leq n\}. \end{aligned}$$

The first factor in the right-hand side is strictly positive, and the left-hand side is at most $C_1(R+\varepsilon_0)^{-n}$. Thus (1.22) also holds for $x = y_0$ (after a suitable change of C_1). Now adding y_0 to $\mathscr{U}_1$ can only help for the inequality in (1.22).

Equation (1.24) holds automatically, because if the inequality holds for $x \in \mathscr{U}_1$ and for $x = y_0$, then it holds for $x \in \mathscr{U}_1 \cup \{y_0\}$. □

From now on we assume that $\mathscr{U}_2$ is a singleton, say $\mathscr{U}_2 = \{y_0\}$, and that this y_0 belongs to $\mathscr{U}_1$ and that x_0 in (1.23) equals y_0.

LEMMA 2. *There exist a constant $C_6 > 0$ and, for all $x \in \mathscr{U}_1$, an $n_1 = n_1(x) < \infty$ such that*

$$\sum_{k=0}^{n} P_x\{X_k = y_0 | T > n\} \geq C_6 n, \qquad n \geq n_1. \tag{2.3}$$

PROOF. For n_0 as in (1.24) and $k \leq n - n_0$ one has

$$\begin{aligned}
&\sum_{l=0}^{n_0} P_x\{X_{k+l} = y_0 | T > n\} \\
&\quad \geq \frac{1}{P_x\{T > n\}} \sum_{u \in \mathscr{U}_1} P^k(x, u) \sum_{l=0}^{n_0} P^l(u, y_0) P_{y_0}\{T > n-k\} \\
&\quad \geq C_3 C_2^{-1} \sum_{u \in \mathscr{U}_1} P^k(x, u) \frac{P_u\{T > n-k\}}{P_x\{T > n\}} \qquad \text{[by (1.23) and (1.24)]} \\
&\quad = C_3 C_2^{-1} P_x\{X_k \in \mathscr{U}_1 | T > n\}.
\end{aligned}$$

It therefore suffices to prove that for some $D_1 > 0$,

$$\sum_{k=0}^{n} P_x\{X_k \in \mathscr{U}_1 | T > n\} \geq D_1 n, \qquad x \in \mathscr{U}_1,\ n \geq n_1. \tag{2.4}$$

Now, let $0 \leq \sigma_0 < \sigma_1 < \cdots$ be the successive times at which $X_\bullet$ visits $\mathscr{U}_1$ and define

$$\lambda(n) = \max\{j\colon\ \sigma_j \leq n\}.$$

We shall take $X_0 = x \in \mathscr{U}_1$ so that $\sigma_0 = 0$ and $\lambda(n)$ will be well defined; $\lambda(n) + 1$ will be the number of visits by $X_\bullet$ to $\mathscr{U}_1$ during $[0, n]$. We define

$$\mathscr{F}_k = \sigma\text{-field generated by } X_0, \ldots, X_k.$$

Then, conditionally on $\mathscr{F}_{\sigma_j}$, $\sigma_{j+1} - \sigma_j$ has a defective distribution, which satisfies on the event $\{\sigma_j < \infty,\ X_{\sigma_j} = y\}$ with $y \in \mathscr{U}_1$,

$$\begin{aligned} P_x\{r < \sigma_{j+1} - \sigma_j < \infty | \mathscr{F}_{\sigma_j}\} \\ &= P_y\{r < \sigma_1 < \infty\} \\ &= P_y\{X_l \notin \mathscr{U}_1 \text{ for } 1 \le l \le r, \text{ but } X_k \text{ returns to } \mathscr{U}_1 \text{ for some } k > r\} \\ &\le P_y\{T > r, \text{ but } X_l \notin \mathscr{U}_1 \text{ for } 1 \le l \le r\} \le C_1(R + \varepsilon_0)^{-r} \end{aligned} \tag{2.5}$$

[by (1.22)]. Consequently, for any $\theta \ge 0$ with $e^\theta < R + \varepsilon_0$, on the event $\{\sigma_j < \infty\}$,

$$\begin{aligned} \sum_{r=1}^{\infty} e^{\theta r} P_x\{\sigma_{j+1} - \sigma_j = r | \mathscr{F}_{\sigma_j}\} \\ &\le e^\theta + (e^\theta - 1) \sum_{r=1}^{\infty} e^{\theta r} P_x\{r < \sigma_{j+1} - \sigma_j < \infty | \mathscr{F}_{\sigma_j}\} \\ &\le D_2(\theta) \end{aligned}$$

for some $1 < D_2(\theta) < \infty$. It follows that

$$\begin{aligned} P_x\{j\text{th return to } \mathscr{U}_1 \text{ occurs at time } r\} \\ &= P_x\{\sigma_j = r\} \\ &\le e^{-r\theta} E_x\left\{ \exp\left(\sum_{i=0}^{j-1} \theta(\sigma_{i+1} - \sigma_i) \right);\ \sigma_j < \infty \right\} \\ &\le e^{-r\theta} [D_2(\theta)]^j. \end{aligned} \tag{2.6}$$

We use this to prove that for some $\alpha > 0$ (independent of $x \in \mathscr{U}_1$) and n_1,

$$P_x\{\lambda(n) \le \alpha n | T > n\} \le \tfrac{1}{2}, \qquad x \in \mathscr{U}_1,\ n \ge n_1, \tag{2.7}$$

which in turn will imply

$$\begin{aligned} \sum_{k=0}^{n} P_x\{X_k \in \mathscr{U}_1 | T > n\} \\ &= E_x\{\text{number of visits by } X_\bullet \text{ during } [0, n] \text{ to } \mathscr{U}_1 | T > n\} \\ &\ge E_x\{\lambda(n) | T > n\} \ge \alpha n P_x\{\lambda(n) > \alpha n | T > n\} \ge \tfrac{1}{2} \alpha n. \end{aligned}$$

Thus (2.4), and hence the lemma, will follow from (2.7). However, for $x \in \mathscr{U}_1$, by a standard last exit decomposition,

$$
\begin{aligned}
&P_x\{\lambda(n) \le \alpha n,\ T > n\} \\
&\quad = \sum_{0 \le j \le \alpha n} \sum_{r=0}^{n} \sum_{y \in \mathscr{U}_1} P_x\{\lambda(n) = j,\ \sigma_{\lambda(n)} = r,\ X_r = y,\ T > n\} \\
&\quad = \sum_{0 \le j \le \alpha n} \sum_{r=0}^{n} \sum_{y \in \mathscr{U}_1} P_x\{\sigma_j = r,\ X_r = y\} \\
&\qquad\qquad \times P_y\{T > n - r,\ \text{but } X_l \notin \mathscr{U}_1 \text{ for } 1 \le l \le n - r\} \\
&\quad \le \sum_{0 \le j \le \alpha n} \sum_{r=0}^{n} P_x\{\sigma_j = r\} C_1 (R + \varepsilon_0)^{-n+r} \quad \text{[by (1.22)]} \\
&\quad \le C_1 \sum_{0 \le j \le \alpha n} \sum_{r=0}^{n} e^{-r\theta} [D_2(\theta)]^j (R + \varepsilon_0)^{-n+r} \\
&\quad \le D_3(\theta) [D_2(\theta)]^{\alpha n} (R + \varepsilon_0)^{-n} \sum_{r=0}^{n} [(R + \varepsilon_0) e^{-\theta}]^r.
\end{aligned} \tag{2.8}
$$

Now choose $\theta > 0$ such that $e^\theta = R + \frac{7}{8}\varepsilon_0$, and then $\alpha > 0$ so small that

$$[D_2(\theta)]^\alpha e^{-\theta} \le \left(R + \tfrac{3}{4}\varepsilon_0\right)^{-1}.$$

We then find that the left-hand side of (2.7) is at most

$$D_4(\theta)\left(R + \tfrac{3}{4}\varepsilon_0\right)^{-n} [P_x\{T > n\}]^{-1}. \tag{2.9}$$

On the other hand, for $x \in S_0$ there exists an $n_1 < \infty$ such that

$$P_x\{T > n\} \ge \widehat{P}^n(x, x) \ge \left(R + \frac{\varepsilon_0}{2}\right)^{-n}, \qquad n \ge n_1, \tag{2.10}$$

by virtue of (1.13) (recall that we took $p = 1$). Equations (2.8)–(2.10) show that if n_1 is chosen large enough, also

$$P_x\{\lambda(n) \le \alpha n | T > n\} \le D_4(\theta) \left(\frac{R + \varepsilon_0/2}{R + 3\varepsilon_0/4}\right)^n \le \frac{1}{2}, \qquad n \ge n_1, \tag{2.11}$$

as desired. □

The preceding lemma showed that there is a reasonable chance for the "average" k that $X_k = y_0$. We now show that $\{X_k \in \mathscr{U}_1$ for some k "close to n"$\}$ has a high probability (all this conditioned on $\{T > n\}$).

LEMMA 3. *For every $\eta > 0$ and $x \in \mathscr{U}_1$ there exist $m_0 = m_0(x, \eta)$ and $n_2 = n_2(x, \eta) < \infty$ such that*

$$P_x\{X_l \notin \mathscr{U}_1 \textit{ for all } n - m_0 \le l \le n | T > n\} \le \eta, \qquad n \ge n_2. \tag{2.12}$$

PROOF. Take $n_1(x)$ as in (2.10). Then for any $x \in \mathscr{U}_1$, $y \in S_0$ and $p \ge n_1(x)$

$$P_x\{T > n\} \ge P^p(x, x) P_x\{T > n - p\} \\ \ge \left(R + \frac{\varepsilon_0}{2}\right)^{-p} P_x\{T > n - p\} \tag{2.13}$$

and for $m \ge n_1(x)$,

$$\begin{aligned} &P_x\{T > n,\ \text{last visit to } \mathscr{U}_1 \text{ by } X_\bullet \text{ occurs before time } n - m\} \\ &\quad = \sum_{r=0}^{n-m-1} P_x\{X_r \in \mathscr{U}_1, X_l \notin \mathscr{U}_1 \text{ for } r < l \le n, T > n\} \\ &\quad = \sum_{r=0}^{n-m-1} \sum_{y \in \mathscr{U}_1} P^r(x, y) P_y\{X_l \notin \mathscr{U}_1 \text{ for } 1 \le l \le n - r, T > n - r\} \\ &\quad \le \sum_{p=m+1}^{n} \sum_{y \in \mathscr{U}_1} P^{n-p}(x, y) C_1 (R + \varepsilon_0)^{-p} \quad [\text{by } (1.22)] \\ &\quad \le C_1 \sum_{p=m+1}^{n} P_x\{T > n - p\}(R + \varepsilon_0)^{-p} \\ &\qquad\qquad (\text{because } X_{n-p} = y \in \mathscr{U}_1 \text{ implies } T > n - p) \\ &\quad \le D_1 \left(\frac{R + (\varepsilon_0/2)}{R + \varepsilon_0}\right)^m P_x\{T > n\} \quad [\text{by } (2.13)]. \end{aligned}$$

Thus, for any $\eta > 0$, we can fix $m_0 = m_0(x, \eta)$ such that for all large n, say $n \ge n_2(x, \eta)$, (2.12) holds. □

We now combine Lemmas 2 and 3 to get our first real recurrence statement.

LEMMA 4. *There exist constants $C_7 > 0$ and $n_3 < \infty$ such that for $n \ge n_3$,*

$$\sum_{k=0}^{n} P_{y_0}\{X_k = y_0 \text{ and } X_n = y_0 | T > n\} \ge C_7 n. \tag{2.14}$$

PROOF. We start from (2.3) with $x = y_0 \in \mathscr{U}_1$. We now take $\eta = C_6/4$ and subtract from (2.3) the following consequence of (2.12):

$$\sum_{k=0}^{n} P_{y_0}\{X_l \notin \mathscr{U}_1 \text{ for all } n - m_0 \le l \le n | T > n\} \\ \le \eta(n+1) \le \tfrac{1}{2} C_6 n, \qquad n \ge n_2(y_0, \eta) + 2$$

$[m_0 = m_0(y_0, \eta)]$. We obtain that for $n \ge n_1(y_0) \vee (n_2(y_0, \eta) + 2)$,

$$\sum_{k=0}^{n} P_{y_0}\{X_k = y_0 \text{ and } X_l \in \mathscr{U}_1 \text{ for some } n - m_0 \le l \le n | T > n\} \\ \ge \tfrac{1}{2} C_6 n. \tag{2.15}$$

The value of X_n is still unspecified in (2.15) and we must now make X_n itself equal to y_0. First we note that there exist an r and a $D_1 > 0$ so that

$$P^q(y, y_0) \geq D_1 \quad \text{for all } y \in \mathscr{U}_1 \text{ and } r \leq q \leq r + m_0. \tag{2.16}$$

This is so because, by (1.13), there exist an s and a $D_2 > 0$ such that

$$P^q(y_0, y_0) \geq D_2 \quad \text{for all } q \text{ with } s \leq q \leq s + n_0 + m_0.$$

Thus for $y \in \mathscr{U}_1$, $s + n_0 \leq q \leq s + n_0 + m_0$,

$$\begin{aligned} P^q(y, y_0) &\geq \sum_{l=0}^{n_0} P_y\{\text{ first visit to } y_0 \text{ occurs at time } l\} P^{q-l}(y_0, y_0) \\ &\geq D_2 \sum_{l=0}^{n_0} P_y\{\text{ first visit to } y_0 \text{ occurs at time } l\} \geq D_2 C_3. \end{aligned}$$

This is (2.16) for $r = s + n_0$ and $D_1 = D_2 C_3$. Finally, (2.15) and (2.16) show that

$$\begin{aligned} &\sum_{k=0}^{n} P_{y_0}\{X_k = y_0,\ X_{n+r} = y_0\} \\ &\quad \geq \sum_{k=0}^{n} \sum_{l=n-m_0}^{n} \sum_{y \in \mathscr{U}_1} P_{y_0}\{X_k = y_0,\ \text{first visit to } \mathscr{U}_1 \\ &\qquad\qquad \text{at or after time } n - m_0 \\ &\qquad\qquad \text{is at time } l \text{ and } X_l = y\} P^{n+r-l}(y, y_0) \\ &\quad \geq D_1 \sum_{k=0}^{n} P_{y_0}\{X_k = y_0,\ X_l \in \mathscr{U}_1 \text{ for some } n - m_0 \leq l \leq n\} \\ &\quad \geq \tfrac{1}{2} D_1 C_6 n P_{y_0}\{T > n\} \geq \tfrac{1}{2} D_1 C_6 n P_{y_0}\{T > n + r\}. \end{aligned} \tag{2.17}$$

Short of replacing $n + r$ by n, this is (2.14) with $C_7 = \frac{1}{4} D_1 C_6$ and $n_3 = (n_1 \vee n_2 \vee r) + 2$. □

LEMMA 5. *The transition probability matrix* $\widehat{P}$ *is R-positive-recurrent and there exist strictly positive functions* f, μ: $S_0 \to (0, \infty)$ *which satisfy* (1.17)–(1.20). *Up to multiplicative constants*, f *and* μ *are unique.*

PROOF. Define

$$g_n = P_{y_0}\{X_n = y_0,\ X_r \neq y_0 \text{ for } 1 \leq r \leq n - 1\}$$

and

$$L = \sum_{n=1}^{\infty} g_n R^n.$$

Then, by Vere-Jones [(1967), Theorem C], $L \le 1$, and $\widehat{P}$ is not R-recurrent is equivalent to

$$L < 1. \tag{2.18}$$

We begin by proving that (2.18) is not consistent with (2.14). To see this, note that (2.14) implies

$$\begin{aligned} \sum_{k=0}^{n} \widehat{P}^k(y_0, y_0)\widehat{P}^{n-k}(y_0, y_0) &\ge C_7 n P_{y_0}\{T > n\} \\ &\ge C_7 n \widehat{P}^n(y_0, y_0), \qquad n \ge n_3 \end{aligned} \tag{2.19}$$

(again because $X_n = y_0 \in S_0$ implies $T > n$). Next introduce i.i.d. random variables $Y, Y_1, Y_2, \dots$ with the distribution

$$\begin{aligned} P\{Y = n\} &= g_n^* := g_n R^n, \qquad 1 \le n < \infty, \\ P\{Y = \infty\} &= 1 - L. \end{aligned} \tag{2.20}$$

[This is really a distribution by the definition of L and (2.18).] Under (2.18) the mass of this distribution on ∞ is strictly positive. Now set

$$T_l = \sum_{j=1}^{l} Y_j$$

and observe that

$$\begin{aligned} P_{y_0}\{X_k = y_0 | X_n = y_0\} &= \frac{\widehat{P}^k(y_0, y_0)\widehat{P}^{n-k}(y_0, y_0)}{\widehat{P}^n(y_0, y_0)} \\ &= \frac{R^k \widehat{P}^k(y_0, y_0) R^{n-k} \widehat{P}^{n-k}(y_0, y_0)}{R^n \widehat{P}^n(y_0, y_0)} \\ &= \frac{P\{k \text{ equals some } T_l\} P\{n-k \text{ equals some } T_l\}}{P\{n \text{ equals some } T_l\}} \\ &= P\{k \text{ equals some } T_l | n \text{ equals some } T_l\}. \end{aligned}$$

Therefore, (2.19) shows that for $n \ge n_3$,

$$\begin{aligned} C_7 n &\le \sum_{k=0}^{n} \frac{\widehat{P}^k(y_0, y_0)\widehat{P}^{n-k}(y_0, y_0)}{\widehat{P}^n(y_0, y_0)} \\ &= E\{\text{number of } T_l \text{ in } [0, n] \big| n \text{ equals some } T_l\} \\ &\le \frac{1}{2} C_7 n + (n+1) P\Big\{\text{number of } T_l \text{ in } [0, n] \text{ exceeds } \frac{1}{2} C_7 n \Big| \\ &\qquad\qquad n \text{ equals some } T_l\Big\}. \end{aligned}$$

However, for any $\eta > 0$,

$$P\{n \text{ equals some } T_l\} = R^n \widehat{P}^n(y_0, y_0) \geq (1-\eta)^n$$

for large n, by virtue of (1.13). Therefore

$$\begin{aligned} &P\{\text{number of } T_l \text{ in } [0,n] \text{ exceeds } \tfrac{1}{2}C_7 n\} \\ &\qquad \geq \tfrac{1}{4}C_7 P\{n \text{ equals some } T_l\} \geq \tfrac{1}{4}C_7(1-\eta)^n \end{aligned} \tag{2.21}$$

for all large n. This, however, is impossible if

$$L^{C_7/2} < 1 - \eta,$$

because

$$\begin{aligned} &P\{\text{number of } T_l \text{ in } [0,n] \text{ exceeds } \tfrac{1}{2}C_7 n\} \\ &\qquad \leq P\{Y_j < \infty \text{ for } j \leq \tfrac{1}{2}C_7 n + 1\} \leq L^{C_7 n/2}. \end{aligned}$$

This is the required contradiction and we conclude that $L = 1$ and that $\widehat{P}$ is R-recurrent.

Next we show that $\widehat{P}$ is R-positive-recurrent. Multiply (2.19) by $s^n R^n$ and sum over n. If we write

$$U(s) = \sum_{n=0}^{\infty} \widehat{P}^n(y_0, y_0) s^n R^n,$$

we obtain

$$U^2(s) \geq C_7 s U'(s) - C_7(n_3+1)^2 R^{n_3}, \qquad 0 \leq s < 1 \tag{2.22}$$

[the $C_7(n_3+1)^2 R^{n_3}$ is needed because (2.19) holds only for $n \geq n_3$]. Now the R-recurrence of $\widehat{P}$ simply says that

$$U(1) = \sum_0^{\infty} R^n \widehat{P}^n(y_0, y_0) = \infty.$$

Therefore, for $\frac{1}{2} \leq t < 1$ and some $D_1 < \infty$,

$$\frac{1}{U(t)} = \frac{1}{U(t)} - \lim_{s \uparrow 1} \frac{1}{U(s)} = \int_t^1 \frac{U'(s)}{U^2(s)}\, ds \leq D_1(1-t).$$

Consequently,

$$U(t) \geq \frac{1}{D_1(1-t)} \quad \text{as} \quad t \uparrow 1$$

and necessarily

$$\limsup_{n \to \infty} R^n \widehat{P}^n(y_0, y_0) > 0,$$

which gives R-positive-recurrence.

The existence of f and μ which satisfy (1.17)–(1.20) and their uniqueness up to multiplicative constants is now guaranteed by Vere-Jones [(1967), Theorem 4.1 and Corollary on page 375]. □

LEMMA 6. *The f and μ of Lemma 5 satisfy (1.21), (1.26) and (1.27).*

PROOF. Assume that (1.21) fails, so that

$$\sum_{y\in S_0} \mu(y) = \infty.$$

Then, by (1.20) and Fatou's lemma

$$\liminf_{n\to\infty} R^n P_{y_0}\{T > n\} = \liminf_{n\to\infty} \sum_{y\in S_0} R^n P_{y_0}\{X_n = y\} \geq f(y_0) \sum_{y\in S_0} \mu(y) = \infty.$$

This, however, is not posible, since (2.14) implies that for $n \geq n_3$,

$$\begin{aligned} C_7 \frac{n}{n+1} &\leq \frac{1}{n+1} \sum_{k=0}^{n} P_{y_0}\{X_k = y_0,\ X_n = y_0 | T > n\} \\ &\leq \frac{P^n(y_0, y_0)}{P_{y_0}(T > n)} = \frac{R^n P^n(y_0, y_0)}{R^n P_{y_0}\{T > n\}} \\ &\sim \frac{f(y_0)\mu(y_0)}{R^n P_{y_0}\{T > n\}} \quad \text{[by (1.20)]}. \end{aligned}$$

Thus (1.21) must hold.

Equation (1.26) now follows from the dominated convergence theorem. Indeed, by (1.18),

$$\mu(x)\widehat{P}^n(x, y) \leq \mu\widehat{P}^n(y) = R^{-n}\mu(y)$$

so that

$$R^n \widehat{P}^n(x, y) \leq \frac{\mu(y)}{\mu(x)}.$$

Thus, by (1.20),

$$\lim_{n\to\infty} R^n P_x\{T > n\} = \lim_{n\to\infty} \sum_{y\in S_0} R^n \widehat{P}^n(x, y) = f(x) \sum_{y\in S_0} \mu(y).$$

Finally, (1.27) is immediate from (1.26) and (1.20). □

LEMMA 7. *The measure $\widetilde{\mu}$, as defined in (1.25), is the unique minimal normalized quasi-stationary distribution.*

PROOF. Clearly $\widetilde{\mu}(S_0) = 1$. Moreover, by (1.18),

$$\widetilde{\mu}\widehat{P}^n = R^{-n}\widetilde{\mu}\ ,$$

so that μ is a normalized quasi-stationary distribution.

To show the minimality of $\widetilde{\mu}$, observe that for any normalized quasi-stationary distribution ν, (1.6) implies

$$r(n+m) = r(n)r(m) \quad \text{and} \quad r(n) = [r(1)]^n$$

and hence, by virtue of (1.7),

$$P_\nu\{T > n\} = [P_\nu\{T > 1\}]^n,$$

$$E_\nu T = \sum_{n=0}^{\infty} P_\nu\{T > n\} = [1 - P_\nu\{T > 1\}]^{-1}.$$

Thus, in the set of normalized quasi-stationary distributions the expected absorption times are ordered in the same way as $P_\nu\{T > 1\}$, and a normalized quasi-stationary distribution $\tilde{\mu}$ is minimal if and only if $P_\nu\{T > 1\}$ is minimized for $\nu = \tilde{\mu}$. However, (1.6) implies that $r(1) = \widehat{P}_\nu\{T > 1\}$ [cf. (1.7)], so that it suffices for the lemma to show that

$$r(1) \geq R^{-1} \quad \text{for any normalized quasi-stationary distribution } \nu. \tag{2.23}$$

However, (2.23) follows from (1.13) and

$$\widehat{P}^n(x, y) \leq [r(1)]^n \frac{\nu(y)}{\nu(x)}$$

[cf. (1.6) and (1.7)].

Finally, there is only one minimal normalized quasi-stationary distribution because the solution of (1.18) is unique, up to a multiplicative constant [see Vere-Jones (1967), Theorem 4.1]. □

LEMMA 8. *Inequality* (1.28) *holds.*

PROOF. We saw in Lemma 1 that we can take x_0 equal to any state and that this may be the same state as y_0. We therefore only have to prove

$$g_n = P_{y_0}\{X_n = y_0, \text{ but } X_r \neq y_0 \text{ for } 1 \leq r \leq n-1\} \leq D_1(R + \eta_0)^{-n} \tag{2.24}$$

for some $D_1 < \infty$, $\eta_0 > 0$, under the assumption that (1.23) holds with x_0 replaced by y_0. Now introduce a Markov chain X_n^* with state space S_0 and (honest) transition probability matrix

$$P^*(x, y) = \frac{RP(x, y)f(y)}{f(x)}, \qquad x, y \in S_0.$$

Analogously to Section 1, write P_x^* for the distribution of X^*-paths conditioned on $X_0^* = x$. It is well known and easy to check that the R-positive-recurrence of $X_\bullet$ implies that $X_\bullet^*$ is positive-recurrent.

We shall first prove that for some $\alpha > 0$, $\eta_1 > 0$ and $D_2 < \infty$,

$$P_{y_0}^*\{X_r^* \in \mathscr{U}_1 \text{ for fewer than } \alpha n \text{ values of } r \text{ in } [0, n]\} \leq D_2(1 - \eta_1)^n. \tag{2.25}$$

Note the similarity of (2.25) to (2.11). The principal difference is that we now use the measure $P_{y_0}^*$, instead of the measure P_x, conditioned on $\{T > n\}$, of (2.11). Even though this will not be needed in the sequel, we point out that there exists a general relationship between the measures P_x conditioned on $\{T > n\}$ and P^*_x. The latter is a limit of the former in the following sense:

$$\begin{aligned}\lim_{n\to\infty} & P_x\{X_1 = x_1, \dots, X_k = x_k | T > n\} \\ &= \lim_{n\to\infty} \frac{P_{x_k}\{T > n-k\}}{P_x\{T > n\}} P_x\{X_1 = x_1, \dots, X_k = x_k\} \\ &= R^k \frac{f(x_k)}{f(x)} P_x\{X_1 = x_1, \dots, X_k = x_k\} \quad \text{[by (1.26)]} \\ &= R^k \frac{f(x_k)}{f(x_{k-1})} \cdots \frac{f(x_1)}{f(x)} P_x\{X_1 = x_1, \dots, X_k = x\} \\ &= P^*_x\{X^*_1 = x_1, \dots, X^*_k = x_k\}.\end{aligned}$$

Returning to the proof of (2.25), we have

$$\begin{aligned}&P^*_{y_0}\{X^*_r \in \mathscr{U}_1 \text{ for fewer than } \alpha n \text{ values of } r \text{ in } [0, n]\} \\ &\quad = \sum_{m=n+1}^{\infty} \sum_{y\in\mathscr{U}_1} \frac{R^m f(y)}{f(y_0)} P_{y_0}\{\lfloor \alpha n + 1\rfloor\text{th visit by } X_\bullet \text{ to } \mathscr{U}_1 \\ &\qquad\qquad \text{occurs at time } m \text{ and } X_m = y\}.\end{aligned} \tag{2.26}$$

Moreover, by (1.26) and (1.23), we have for some $D_3 < \infty$,

$$\frac{f(y)}{f(y_0)} = \lim_{n\to\infty} \frac{P_y\{T > n\}}{P_{y_0}\{T > n\}} \le D_3 \quad \text{for all } y \in \mathscr{U}_1. \tag{2.27}$$

Therefore, the left-hand side of (2.26) is (for large n) at most

$$\begin{aligned}2D_3 &\sum_{m=n+1}^{\infty} \sum_{y\in\mathscr{U}_1} R^m P_{y_0}\{\lfloor \alpha n + 1\rfloor\text{-th visit by } X_\bullet \text{ to } \mathscr{U}_1 \text{ occurs at time } m, \\ &\qquad X_m = y, T > m\} \\ &\qquad\qquad (\text{again because } X_m = y \in \mathscr{U}_1 \text{ implies } T > m) \\ &\le D_4 \sum_{m=n+1}^{\infty} \sum_{y\in\mathscr{U}_1} P_{y_0}\{\lfloor \alpha n + 1\rfloor\text{th visit by } X_\bullet \text{ to } \mathscr{U}_1 \text{ occurs at time } m, \\ &\qquad X_m = y | T > m\} \quad \text{[by (1.26)]} \\ &\le D_4 \sum_{m=n+1}^{\infty} P_{y_0}\{X_r \in \mathscr{U}_1 \text{ for fewer than } (\alpha n + 2) \text{ values} \\ &\qquad \text{of } r \text{ in } [0, m] | T > m\} \\ &\le D_5 \sum_{m=n+1}^{\infty} \left(\frac{R + \varepsilon_0/2}{R + 3\varepsilon_0/4}\right)^m \quad \text{[by (2.11)].}\end{aligned}$$

This proves (2.25) with

$$\eta_1 = \frac{\varepsilon_0}{4R + 3\varepsilon_0}.$$

The second ingredient for the proof is the following simple estimate, which holds uniformly for $y \in \mathscr{U}_1$ and $k \geq 0$:

$$\begin{aligned}
&P^*_{y_0}\{X^*_r = y_0 \text{ for some } k \leq r \leq k + n_0 | X^*_k = y\} \\
&\quad\geq \frac{1}{n_0+1} E^*_{y_0}\{\text{number of } r \text{ in } [k, k+n_0] \text{ with } X^*_r = y_0 | X^*_k = y\} \\
&\quad= \frac{1}{n_0+1} \sum_{j=0}^{n_0} (P^*)^j (y, y_0) \\
&\quad\geq \frac{f(y_0)}{n_0+1}\Big[\sup_{y\in\mathscr{U}_1} f(y)\Big]^{-1} \sum_{j=0}^{n_0} P^j(y, y_0) \qquad (2.28)\\
&\quad\geq C_3 \frac{f(y_0)}{n_0+1}\Big[\sup_{y\in\mathscr{U}_1} f(y)\Big]^{-1} \quad [\text{by } (1.24)] \\
&\quad\geq D_6 > 0 \quad [\text{by } (2.27)]
\end{aligned}$$

for some constant $D_6 > 0$. If $0 \leq \sigma^*_0 < \sigma^*_1 < \cdots$ are the times of the successive visits by $X^*_\bullet$ to $\mathscr{U}_1$ and

$$\mathscr{F}^*_k = \sigma\text{-field generated by } X^*_0, \dots, X^*_k,$$

then (2.28) implies that

$$\begin{aligned}
&P^*_{y_0}\{X^*_r \text{ equals } y_0 \text{ for some } \sigma^*_j \leq r \leq \sigma^*_{j+n_0} | \mathscr{F}^*_{\sigma_j}\} \\
&\quad\geq P^*_{y_0}\{X^*_r \text{ equals } y_0 \text{ for some } \sigma^*_j \leq r \leq \sigma^*_j + n_0 | \mathscr{F}^*_{\sigma_j}\} \\
&\quad\geq D_6.
\end{aligned}$$

This in turn shows that

$$P^*_{y_0}\{X^*_r \neq y_0 \text{ for all } 1 \leq r \leq \sigma^*_l\} \leq (1 - D_6)^{(l-1)/(n_0+1)}. \qquad (2.29)$$

Finally note that

$$\begin{aligned}
g^*_n &:= P^*_{y_0}\{X^*_n = y_0 \text{ but } X^*_r \neq y_0 \text{ for } 1 \leq r \leq n-1\} \\
&= \sum_{x_1,\dots,x_{n-1}\neq y_0} P^*(y_0, x_1)P^*(x_1, x_2)\cdots P^*(x_{n-1}, y_0) \\
&= R^n \sum_{x_1,\dots,x_{n-1}\neq y_0} P(y_0, x_1)\cdots P(x_{n-1}, y_0) = R^n g_n,
\end{aligned}$$

so that (2.24) is equivalent to

$$g^*_n \leq D_1 \left(\frac{R}{R+\eta_0}\right)^n. \qquad (2.30)$$

This follows for suitable $\eta_0 > 0$, $D_1 < \infty$ from (2.25) and (2.29) by observing that

$$\begin{aligned} g_n^* &\le P_{y_0}^*\{X_r^* \in \mathscr{U}_1 \text{ for fewer than } \alpha n \text{ values of } r \text{ in } [0, n]\} \\ &\quad + P_{y_0}^*\{X_r^* \neq y_0 \text{ up until the } \alpha n\text{th visit of } X^* \text{ to } \mathscr{U}_1\} \\ &\le D_2(1-\eta_1)^n + (1-D_6)^{(\lfloor \alpha n \rfloor - 2)/(n_0+1)}. \end{aligned}$$ □

We have now proven all statements in Theorem 1.

3. Cellular automata. In this section we discuss Markov chains of the type discussed before Theorem 2 in the Introduction and prove Theorem 2. As we shall show, the main reason for the validity of Theorem 2 is that the number of occupied sites in its chains does have a very strong downward drift when this number of occupied sites becomes large. This is made precise by the estimate (3.2) below.

PROOF OF THEOREM 2. We shall occasionally write $\widetilde{X}(n, z)$ and $|\widetilde{X}(n)|$ for $\widetilde{X}_n(z)$ and $|\widetilde{X}_n|$, respectively, for typographical convenience. We remind the reader that the range ρ is introduced in (1.35). Because $\widetilde{X}_{\bullet}$ is strongly subcritical, there exists an $n_4 \ge 1$ such that for all $x \in \widetilde{S}$, $z \in \mathbb{Z}^d$,

$$P_x\{\widetilde{X}(n_4, z) \neq 0\} \le \tfrac{1}{2}(2\rho + 1)^{-d} n_4^{-d}. \tag{3.1}$$

We claim that this implies

$$E_x\{|\widetilde{X}(n_4)|\} \le \tfrac{1}{2}|x| \quad \text{for all } x \in \widetilde{S}. \tag{3.2}$$

To see this, write

$$|\widetilde{X}(n_4)| = \sum_{z \in \mathbb{Z}^d} I[n_4, z], \tag{3.3}$$

where $I[n, z]$ is the indicator function of the event $\{\widetilde{X}(n, z) \neq 0\}$. Now the distribution of $\widetilde{X}(n, z)$ depends on x only through the values of $x(v)$ with $v \in \mathbb{Z}^d$ such that $\|v - z\| \le n\rho$ (where $\| \;\|$ is short for $\| \;\|_2$). For $n = 1$ this is just assumption (1.35), and it easily follows for general n by induction on n. Consequently, if $x(v) = 0$ for all v with $\|v - z\| \le n\rho$, then

$$P_x\{\widetilde{X}(n, z) \neq 0\} = P_{\underline{0}}\{\widetilde{X}(n, z) \neq 0\} = 0, \tag{3.4}$$

because the state $\underline{0}$, with all components equal to 0, is absorbing. It follows that if $v_1, \dots, v_{|x|}$ are the sites in $\mathbb{Z}^d$ with $x(v) \neq 0$, then a.e. $[P_x]$,

$$|\widetilde{X}(n_4)| \le \sum_{j=1}^{|x|} \sum_{z \in \overline{B}(v_j, n_4\rho)} I[n_4, z], \tag{3.5}$$

where

$$\overline{B}(v, r) = \{z \in \mathbb{Z}^d \colon \|z - v\| \le r\}.$$

Taking expectation with respect to P_x in (3.5) gives

$$\begin{aligned} E_x|\widetilde{X}(n_4)| &\le \sum_{j=1}^{|x|} \sum_{z\in\overline{B}(v_j,\, n_4\rho)} P_x\{\widetilde{X}(n_4, z) \neq 0\} \\ &\le |x||\overline{B}(0, n_4\rho)|\tfrac{1}{2}(2\rho+1)^{-d} n_4^{-d} \quad \text{[by (3.1)]} \\ &\le \tfrac{1}{2}|x|. \end{aligned}$$

This establishes (3.2).

Now for any integer $m \ge 1$ and $|x| \le K$,

$$\begin{aligned} &P_x\{|\widetilde{X}(jn_4)| > K \text{ for } 1 \le j \le m,\ T > mn_4\} \\ &\quad\le \frac{1}{K^r} E_x\{|\widetilde{X}(mn_4)|^r;\ |\widetilde{X}(jn_4)| > K,\ 1 \le j \le m\} \\ &\quad= \frac{|x|^r}{K^r} E_x\left\{\prod_{j=0}^{m-1} \frac{|\widetilde{X}((j+1)n_4)|^r}{|\widetilde{X}(jn_4)|^r};\ |\widetilde{X}(jn_4)| > K,\ 1 \le j \le m\right\} \\ &\quad\le \sup_{|x|\le K} E_x\left(\frac{|\widetilde{X}(n_4)|^r}{|x|^r}\right)\left[\sup_{|y|>K} E_y\left(\frac{|\widetilde{X}(n_4)|^r}{|y|^r}\right)\right]^{m-1}. \end{aligned} \tag{3.6}$$

By (3.5),

$$|\widetilde{X}(n_4)| \le |\widetilde{X}(0)|(2\rho+1)^d n_4^d$$

so that

$$\sup_{|x|\le K} E_x\left(\frac{|\widetilde{X}(n_4)|^r}{|x|^r}\right) \le (2\rho+1)^{dr} n_4^{dr}. \tag{3.7}$$

We next fix r such that

$$\left(\frac{2}{3}\right)^{r/n_4} < \gamma \tag{3.8}$$

[where γ is the number appearing in the right-hand side of (1.39)]. We shall show that for all large K,

$$\sup_{|y|>K} E_y\left(\frac{|\widetilde{X}(n_4)|^r}{|y|^r}\right) \le \left(\frac{2}{3}\right)^r. \tag{3.9}$$

Now observe that the absorption time T is the same for the two chains $\{X_n\}$ and $\{\widetilde{X}_n\}$. Also $|X_n| = |\widetilde{X}_n|$. Therefore, if (3.9) holds,

$$\begin{aligned}
\sup_{x\in S, |x|\le K} & P_x\{T > n,\ \text{but } |X_l| > K \text{ for all } 1 \le l \le n\} \\
&= \sup_{x\in \widetilde{S}, |x|\le K} P_x\{T > n,\ \text{but } |\widetilde{X}_l| > K \text{ for all } 1 \le l \le n\} \\
&\le \sup_{x\in \widetilde{S}, |x|\le K} P_x\{|\widetilde{X}(jn_4)| > K \text{ for all } 1 \le j \le \lfloor n/n_4 \rfloor\} \\
&\le (2\rho+1)^{dr} n_4^{dr} \left(\tfrac{2}{3}\right)^{r(\lfloor n/n_4\rfloor - 1)} \quad [\text{by (3.6), (3.7) and (3.9)}] \\
&\le C_5 \gamma^n \quad [\text{by (3.8)}]
\end{aligned}$$

with

$$C_5 = \left(\tfrac{3}{2}\right)^{2r} (2\rho+1)^{dr} n_4^{dr}$$

(which is independent of n, K). Thus (3.9) will imply (1.39) and we now turn to the proof of (3.9).

By (3.3),

$$\begin{aligned}
E_x|\widetilde{X}(n_4)|^r &= \sum_{z_1,\dots,z_r} E_x\{I[n_4, z_1]\cdots I[n_4, z_r]\} \\
&= \sum_{z_1,\dots,z_r} \prod_{i=1}^r E_x I[n_4, z_i] \qquad (3.10)\\
&\quad + \sum_{z_1,\dots,z_r} \left\{E_x\{I[n_4, z_1]\cdots I[n_4, z_r]\} - \prod_{i=1}^r E_x I[n_4, z_i]\right\}
\end{aligned}$$

Also, by (3.2),

$$\sum_{z_1,\dots,z_r} \prod_{i=1}^r E_x I[n_4, z_i] = \left\{E_x \sum_z I[n_4, z]\right\}^r \le \left(\tfrac{1}{2}|x|\right)^r.$$

Moreover,

$$\left|E_x\{I[n_4, z_1]\cdots I[n_4, z_r]\} - \prod_{i=1}^r E_x I[n_4, z_i]\right| \le 1,$$

so that (3.10) shows

$$\begin{aligned}
E_x \frac{|\widetilde{X}(n_4)|^r}{|x|^r} \le 2^{-r} + |x|^{-r} \Big(&\text{number of } r\text{-tuples } z_1, \dots, z_r \text{ with} \qquad (3.11)\\
& E_x\{I[n_4, z_1]\cdots I[n_4, z_r]\} \ne \prod_{i=1}^r E_x I[n_4, z_i]\Big).
\end{aligned}$$

We claim that for any subsets $E_1, \dots, E_r$ of $\mathbb{Z}^d$, which satisfy

$$d(E_i, E_j) = \inf\{\|z' - z''\|: \ z' \in E_i, z'' \in E_j\} > 2n\rho \quad \text{for} \quad i \neq j, \tag{3.12}$$

the families $\{\widetilde{X}(n, z): \ z \in E_i\}$ are independent under P_x. In fact, for $n = 1$ all $\widetilde{X}(1, z)$, $z \in \mathbb{Z}^d$, are independent under P_x by (1.34). For general n our claim follows by induction, by means of (1.34) and (1.35). Indeed if the result is true for $n = 1, \dots, k$ and $E_1, \dots, E_r$ satisfy (3.12) with $n = k + 1$, then for any bounded functions f_i, depending on $\{\widetilde{X}(k + 1, z): \ z \in E_i\}$, $1 \le i \le r$, the Markov property and (1.34) imply that

$$E_x\left\{\prod_1^n f_i\right\} = E_x\left\{E_x\left\{\prod_1^r f_i \middle| \widetilde{X}_k\right\}\right\} = E_x\left\{\prod_1^r E_x\{f_i | \widetilde{X}_k\}\right\} \tag{3.13}$$

(note that the E_i are disjoint). Furthermore, by (1.35), $E_x\{f_i | \widetilde{X}_k\} = g_i$ for some bounded function g_i of the $\widetilde{X}_k(v)$ with $v \in E_i'$, where

$$E_i' = \{v \in \mathbb{Z}^d: \ \|v - z\| \le \rho \text{ for some } z \in E_i\}.$$

The E_i' satisfy (3.12) with $n = k$, so that by the induction hypothesis the g_i are independent under P_x. Thus

$$E_x\left\{\prod_1^r f_i\right\} = E_x\left\{\prod_1^r g_i\right\} = \prod_1^r E_x\{g_i\} = \prod_1^r E_x\{E_x\{f_i | \widetilde{X}_k\}\} = \prod_1^r E_x\{f_i\}.$$

Thus the families $\{\widetilde{X}_{k+1}(z): \ z \in E_i\}$ are independent, as claimed.

The above independence property shows that

$$E_x\{I[n_4, z_1] \cdots I[n_4, z_r]\} = \prod_{i=1}^r E_x I[n_4, z_i]$$

as soon as $\|z_i - z_j\| > 2n_4\rho$ for $i \neq j$. Furthermore, by (3.4), $I[n_4, z_i] = 0$ a.e. $[P_x]$ unless z_i lies within distance $n_4\rho$ from the set

$$\{v \in \mathbb{Z}^d: \ x(v) \neq 0\}. \tag{3.14}$$

Thus

$$E_x\{I[n_4, z_1] \cdots I[n_4, z_r]\} = \prod_{i=1}^r E_x I[n_4, z_i] = 0$$

unless *each* z_i lies within distance $n_4\rho$ from the set (3.14). Since the set in (3.14) has cardinality $|x|$, the number of r-tuples $z_1, \dots, z_r$ which lie entirely within distance $n_4\rho$ of the set (3.14) *and* have $\|z_i - z_j\| \le 2n_4\rho$ for some $i \neq j$ is at most

$$[(2n_4\rho + 1)^d |x|]^{r-1} r^2 (4n_4\rho + 1)^d.$$

This is therefore an upper bound for the number appearing in the right-hand side of (3.11) and, consequently,

$$E_x \frac{|\widetilde{X}(n_4)|^r}{|x^r|} \le 2^{-r} + \frac{1}{|x|} r^2 (4n_4\rho + 1)^{dr}.$$

Equation (3.9) is now immediate for

$$|K| > \frac{r^2(4n_4\rho+1)^{dr}}{(2/3)^r - (1/2)^r}. \qquad \square$$

4. Some examples. We shall now use Theorem 2 to give some specific examples to which Theorem 1 applies. These examples therefore have normalized quasi-stationary distributions and the limit relations (1.26) and (1.27) hold for them. All examples are of the form described before Theorem 2, and we allow only the values 0 or 1 for $\widetilde{X}_n(z)$ [i.e., $\kappa = 2$ in (1.30)]. By the translation invariance (1.33) and the independence property (1.34) the model is completely specified by the probabilities

$$P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0 = x\} \tag{4.1}$$

(where $\mathbf{0}$ denotes the origin in $\mathbb{Z}^d$), and this probability depends only on $x(z)$ for $\|z\| \le \rho$ [by (1.35)]. Note that we have an obvious partial order on $\widetilde{S}_{d,2}$: $x' \ge x''$ if and only if $x'(z) \ge x''(z)$ for all $z \in \mathbb{Z}^d$. We now impose four further conditions on the probabilities in (4.1):

$$P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0 = x\} \text{ is increasing in } x; \tag{4.2}$$

$$P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0 = x\} < 1 \text{ for all } x \in \widetilde{S}_{d,2}; \tag{4.3}$$

(4.4) $P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0 = x\} = 0$ if $x(z) = 0$ for all $z \in \overline{B}(\mathbf{0}, \rho)$, but $P\{\widetilde{X}_1(z_0) = 1 | \widetilde{X}_0 = x\} > 0$ for some z_0 if $x(z) = 1$ for some $z \in \overline{B}(\mathbf{0}, \rho)$;

(4.5) for any disjoint subsets $A_1, \dots, A_s$ of $\overline{B}(\mathbf{0}, \rho) = \{v\colon \|v\| \le \rho\}$, it holds that $P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0(z) = 1$ in $\overline{B}(\mathbf{0}, \rho)$ exactly for the z's in $\bigcup_1^s A_i\} \le 1 - \prod_{i=1}^s P\{\widetilde{X}_1(\mathbf{0}) = 0 | \widetilde{X}_0(z) = 1$ in $\overline{B}(\mathbf{0}, \rho)$ exactly for the z's in $A_i\}$.

Condition (4.5) says that the chain starting with only the sites in $\bigcup_1^s A_i$ occupied lies stochastically below the maximum of s independent chains, the ith of which starts with only the sites in A_i occupied, $1 \le i \le s$. In order to obtain the required irreducibility (1.2) or (1.32), we shall restrict the state space S to

(4.6) $\underline{0} \cup$ collection of states which can be reached by $X_\bullet$ when starting at $X_0 = \delta(\mathbf{0})$

[$\delta(\mathbf{0})$ is the state with only the component at the origin equal to 1 and all others equal to 0]. We shall prove the following result.

PROPOSITION 1. *If* (1.31)–(1.35) *and* (4.2)–(4.5) *hold and if* $\{\widetilde{X}_n\}$ *is strongly subcritical, then* $\{X_n\}$ [*as defined by* (1.36) *and restricted to the space* S *of* (4.6)] *satisfies all hypotheses of Theorem* 1. *It therefore has a minimal normalized quasi-stationary distribution* $\widetilde{\mu}$, *and the limit relations* (1.26) *and* (1.27) *hold as well as the bound* (1.28).

Before proving the proposition, we give some very specific examples.

EXAMPLE A (One-dimensional nearest neighbor chains). These chains have $d = 1$, $\rho = 1$, and $\widetilde{X}_n(z)$ takes on only the values 1 and 0. Thus, given $\widetilde{X}_n$, the distribution of $\widetilde{X}_{n+1}(z)$ depends only on $\widetilde{X}_n(z-1)$, $\widetilde{X}_n(z)$, $\widetilde{X}_n(z+1)$ (remember that $z \in \mathbb{Z}$ now). In fact we shall assume that it depends only on the number of 1's among $\widetilde{X}_n(z-1)$, $\widetilde{X}_n(z)$, $\widetilde{X}_n(z+1)$. The distribution of $\widetilde{X}_{n+1}$ is therefore completely specified by the four probabilities

$$\pi_i := P\{\widetilde{X}_{n+1}(z) = 1 | \widetilde{X}_n\} \text{ on the event}$$
$$\{\text{exactly } i \text{ of the values } \widetilde{X}_n(z-1), \widetilde{X}_n(z), \widetilde{X}_n(z+1) \text{ equal } 1\},$$
$$0 \le i \le 3.$$

In fact, to make $\underline{0}$ absorbing we must take

$$\pi_0 = 0. \tag{4.7}$$

Equations (4.2)–(4.5) will then hold if we take, in addition to (4.7),

$$0 < \pi_1 \le \pi_2 \le \pi_3 < 1, \tag{4.8}$$

$$\pi_2 \le 1 - (1-\pi_1)^2 = 2\pi_1 - \pi_1^2 \tag{4.9}$$

and

$$\pi_3 \le 1 - (1-\pi_1)(1-\pi_2) = \pi_1 + \pi_2 - \pi_1\pi_2. \tag{4.10}$$

Indeed (4.2)–(4.4) hold by virtue of (4.8). As for (4.5), $\overline{B}(0,1) = \{-1, 0, +1\}$ and the only possible disjoint choices for $A_1, \dots, A_s$ are

$$s = 2, \qquad A_1 = \{-1\}, \qquad A_2 = \{0\}, \tag{4.11}$$

or

$$s = 2, \qquad A_1 = \{-1\}, \qquad A_2 = \{0, 1\}, \tag{4.12}$$

or

$$s = 3, \qquad A_1 = \{-1\}, \qquad A_2 = \{0\}, \qquad A_3 = \{1\}, \tag{4.13}$$

or equivalent choices obtained from (4.11)–(4.13) by permuting $\{-1, 0, 1\}$. Condition (4.5) is void for $s = 1$. For the situations in (4.11) and (4.12), (4.5) is guaranteed by (4.9) and (4.10), respectively. For the case in (4.13), (4.5) requires

$$\pi_3 \le 1 - (1-\pi_1)^3,$$

but this is easily seen to be a consequence of (4.9) and (4.10).

We do not have a sharp criterion for this chain to be strongly subcritical. However, this chain lies stochastically below oriented site percolation with $d = 1$, $p = \pi_3$, discussed in the next example. Thus if π_3 is sufficiently small, then $\{\widetilde{X}_n\}$ is strongly subcritical.

EXAMPLE B (The discrete time contact process or oriented percolation). We describe here a discrete time analogue of the contact process. As will be apparent from the description below, this model can also be viewed as percolation on the vertex set $\mathbb{Z}^d \times \{1, 2, \dots\}$ with an oriented edge from $y \times \{n\}$ to $z \times \{n+1\}$ if and only if y and z are adjacent on $\mathbb{Z}^d$. The usual oriented percolation is essentially the same model but with $\mathbb{Z}^d$ replaced by another lattice; we shall not consider that, though. Again we take a nearest neighbor process on $\mathbb{Z}^d$, that is, with $\rho = 1$. The probability in (4.1) depends only on the number of 1's among the $x(v)$, with v in $\overline{B}(\mathbf{0}, 1) \setminus \{\mathbf{0}\} = \{v\colon\ v \text{ adjacent to } \mathbf{0} \text{ on } \mathbb{Z}^d\}$. We do allow $d > 1$ now, though. There are two versions of this model.

(i) *Site version*. This version corresponds to the threshold contact process. We take $P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0 = x\} = p$ if at least one v adjacent on $\mathbb{Z}^d$ to $\mathbf{0}$ has $x(v) = 1$, and $P\{\widetilde{X}_1(\mathbf{0}) = 1 | \widetilde{X}_0 = x\} = 0$ otherwise. Equation (4.2) is obvious and so are (4.3) and (4.4) if $0 < p < 1$. We shall not check (4.5) formally, but this is almost obvious from the following percolation construction. Color all sites of $\mathbb{Z}^d \times \{1, 2, \dots\}$ white (black) with probability p ($1 - p$, respectively), independently of each other. When $\widetilde{X}_0 = x$ is given, define $\widetilde{X}_n$ for $n \geq 1$ recursively by the following rule. If the site at $z \times \{n\}$ is black, then $\widetilde{X}_n(z) = 0$ [i.e., z is vacant at time n]. If $z \times \{n\}$ is colored white and at least one of its neighbors is occupied at time $n - 1$ [i.e., $\widetilde{X}_{n-1}(v) = 1$ for some neighbor v on $\mathbb{Z}^d$ of z], then z is also occupied at time n. If z has no occupied neighbor at time $n - 1$, then z is not occupied at time n, irrespective of the color of $z \times \{n\}$.

It is known [see for instance Durrett (1988), Section 5a] that this chain has a critical probability $p_c = p_c(d, \text{site}) \in (0, 1)$ such that for $p \leq p_c$ the process becomes extinct, that is, (1.4) holds for any starting configuration with finitely many occupied sites. Moreover, if we write $\delta(z_0)$ for the state x with $x(z) = 1$ if and only if $z = z_0$, then

$$P_{\delta(z_0)}\{T > n\} \to 0 \text{ exponentially in } n \tag{4.14}$$

whenever $p < p_c$ [the probability in (4.14) is independent of z_0]. This can be seen by redoing Menshikov's proof for the usual unoriented percolation [as given in Grimmett (1989), Section 3.2] or by using Aizenman and Barsky (1987) together with Hammersley's theorem [Theorem 5.1 in Grimmett (1989); the latter's proof in the oriented case is quite easy]. For a proof of (4.14) in a more complicated situation, see also Bezuidenhout and Grimmett (1991).

Equation (4.14) implies that

$$\text{for } p < p_c,\ \{\widetilde{X}_n\} \text{ is strongly subcritical.} \tag{4.15}$$

To see this note that as in the lines following (3.3), $\widetilde{X}_n(\mathbf{0})$ depends only on $x(v)$ with $\|v\| \leq n$. Furthermore, by the above description, $\widetilde{X}_n(\mathbf{0}) = 1$ occurs if and only if there is a path $v_0 = \mathbf{0}, v_1, v_2, \dots, v_n$ on $\mathbb{Z}^d$ such that $v_k \times \{n-k\}$ is colored white for $0 \le k \le n-1$ and $x(v_n) = 1$. Therefore,

$$(4.16)\qquad \begin{aligned} P_x\{\widetilde{X}_n(\mathbf{0}) = 1\} &\le \sum_{\|v\| \le n} P_{\delta(v)}\{\widetilde{X}_n(\mathbf{0}) = 1\} \\ &\le (2n+1)^d P_{\delta(\mathbf{0})}\{T > n\} \end{aligned}$$

and this tends to 0 exponentially fast in n. Thus for $p < p_c$ all conclusions of Theorem 1 hold for this chain.

(ii) *Bond version*. We now take

$$P\{\widetilde{X}_1(z) = 1 | \widetilde{X}_0 = x\} = 1 - (1-p)^j$$

if z has exactly j occupied neighbors in state x. In terms of independent colorings we can describe this model as follows: put an edge or bond between $v \times \{n\} \in \mathbb{Z}^d \times \{0, 1, \dots\}$ and $z \times \{n+1\}$ if and only if v and z are adjacent. Color all edges independently white (or black) with probability p (or $1-p$, respectively). Then $\widetilde{X}_{n+1}(z) = 1$ if and only if it has a white edge to an occupied neighbor at time n, that is, if $\widetilde{X}_n(v) = 1$ for some v adjacent to z for which the edge between $v \times \{n\}$ and $z \times \{n+1\}$ is white. Again this chain has a critical probability $p_c = p_c(d, \text{bond}) \in (0, 1)$ such that for $p < p_c$ all conclusions of Theorem 1 hold. We skip the details.

REMARK 7. A closely related problem to (1.27) for one-dimensional oriented subcritical percolation is the following question: Is there a limit distribution for the configuration of occupied sites at time n, as seen from the leftmost occupied site, when in the initial state all sites in $\{0, 1, 2, \dots\}$ are occupied? In this case it is not necessary to condition on $\{T > n\}$ because $T = \infty$ almost surely for any initial state with infinitely many occupied sites. This question was investigated by A. Galves, M. Keane and I. Meilijson (private communication). Our results do not apply directly to this problem, but we hope that they will nevertheless be useful.

REMARK 8. For the discrete time contact process it is almost trivial to verify (1.23) if we take $\mathscr{U}_1 = \{x\colon |x| \le K\}$ for some K and $x_0 = \delta(\mathbf{0})$. Indeed, for any $x \in \mathscr{U}_1$,

$$(4.17)\qquad \begin{aligned} &P_x\{X_n \text{ is not absorbed by time } n\} \\ &\quad = P_x\{\text{one of the } |x| \text{ initial particles survives until time } n\} \\ &\quad \le |x| P_{\delta(\mathbf{0})}\{T > n\} \le K P_{\delta(\mathbf{0})}\{T > n\}. \end{aligned}$$

The next lemma shows how to generalize this estimate under (4.2)–(4.5).

LEMMA 9. *Let* $\{\{\widetilde{X}_n^v\}_{n \ge 0}\colon v \in \mathbb{Z}^d\}$ *be a family of independent Markov chains, each of which has the same transition probabilities as* $\{\widetilde{X}_n\}$, *but* $\widetilde{X}_{\bullet}^v$ *has initial state* $\widetilde{X}_0^v = \delta(v)$. *If* (1.31)–(1.35), (4.2) *and* (4.5) *hold, then* $\{\widetilde{X}_n\}$, *starting from* $\widetilde{X}_0 = x$, *is stochastically smaller than the process* $\{Y_n\}$ *defined by*

$$Y_n(z) = \max\{\widetilde{X}^v_n(z)\colon\ x(v) = 1\}. \tag{4.18}$$

In particular, for any $A \subset \mathbb{Z}^d$,

$$\begin{aligned} P_x\{\widetilde{X}_n(z) = 0 \textit{ for all } z \in A\} &\geq P\{Y_n(z) = 0 \textit{ for all } z \in A\} \\ &= \prod_{\substack{v \textit{ with} \\ x(v)=1}} P\{\widetilde{X}^v_n(z) = 0 \textit{ for all } z \in A\}. \end{aligned} \tag{4.19}$$

PROOF. Equation (4.19) is proven by induction on n, together with a coupling of $\widetilde{X}_\bullet$ and all the $\widetilde{X}^v_\bullet$. For $n = 1$, (4.19) is immediate from (4.5) with $A_i = \{v_i\}$, if $v_1, \dots, v_s$ are the sites for which $x(v) = 1$. Indeed, by (1.34),

$$P_x\{\widetilde{X}_1(z) = 0 \text{ for all } z \in A\} = \prod_{z \in A} P_x\{\widetilde{X}_1(z) = 0\}.$$

Moreover, by (1.35), $P_x\{\widetilde{X}_1(z) = 0\}$ depends only on which v's in $\overline{B}(z, 1)$ have $x(v) = 1$ and, by (4.5),

$$\begin{aligned} P_x\{\widetilde{X}_1(z) = 0\} &\geq \prod_{v_i \in \overline{B}(z,1)} P_{\delta(v_i)}\{\widetilde{X}_1(z) = 0\} \\ &\geq \prod_{i=1}^{s} P_{\delta(v_i)}\{\widetilde{X}_1(z) = 0\} = P\{Y_1(z) = 0\}. \end{aligned} \tag{4.20}$$

Now the $Y_1(z)$, $z \in \mathbb{Z}^d$, are also independent when $\widetilde{X}_0 = x$ is given [by virtue of (1.34)]. Thus (4.20) gives (4.19) for $n = 1$. It is now easy to couple all the $\widetilde{X}^v_1$, $v \in \mathbb{Z}^d$, and $\widetilde{X}_1$. One merely chooses the $\widetilde{X}^v_1$, $v \in \mathbb{Z}^d$, independently of each other, with their prescribed distribution. These $\widetilde{X}^v_1$ determine Y_1. One now takes $\widetilde{X}_1(z) = 0$ for all z with $Y_1(z) = 0$ and takes $\widetilde{X}_1(z) = 0$ (respectively, 1) with probability

$$\frac{P_x\{\widetilde{X}_1(z) = 0\} - P_x\{Y_1(z) = 0\}}{1 - P_x\{Y_1(z) = 0\}}$$

(respectively,

$$\left. \frac{P_x\{\widetilde{X}_1(z) = 1\}}{1 - P_x\{Y_1(z) = 0\}} \right)$$

when $Y_1(z) = 1$. These choices of the $\widetilde{X}_1(z)$, $z \in \mathbb{Z}^d$, are conditionally independent (given x and all the $\widetilde{X}^v_1$, $v \in \mathbb{Z}^d$). One easily checks that the $\widetilde{X}_1(z)$, $z \in \mathbb{Z}^d$, are independent under this construction and, therefore, have the correct distribution. Moreover,

$$\widetilde{X}_1(z) \leq Y_1(z), \qquad z \in \mathbb{Z}^d.$$

Now assume that (4.19) has been proven for $n = 1, \dots, k$ and that we have also constructed joint versions of all $\widetilde{X}_n$, $\widetilde{X}_n^v$, $v \in \mathbb{Z}^d$, $0 \le n \le k$, such that

$$\widetilde{X}_n(z) \le Y_n(z), \qquad z \in \mathbb{Z}^d,\ 0 \le n \le k. \tag{4.21}$$

Let $C_k = \{w_1, \dots, w_r\}$ be the collection of sites for which $\widetilde{X}_k(w) = 1$. A fortiori $Y_k(w_i) = 1$ for $1 \le i \le r$ and hence we can choose a v_i such that $x(v_i) = 1$, $\widetilde{X}_k^{v_i}(w_i) = 1$. The $v_i, 1 \le i \le r$, are not, in general, distinct. Call $w_i \sim w_j$ if $v_i = v_j$. Then C_k breaks up into disjoint equivalence classes, say $A_1, \dots, A_s$. Thus w_i and w_j belong to the same subclass if and only if $v_i = v_j$. We now apply (4.5) with these A_i or rather with $A_i \cap \overline{B}(z, \rho)$ for fixed z. Specifically, we first use

$$P_x\{\widetilde{X}_{k+1}(z) = 0 \text{ for all } z \text{ in } A | \widetilde{X}_k\} = \prod_{z \in A} P_{\widetilde{X}_k}\{\widetilde{X}_1(z) = 0\}.$$

Then, by (4.5), for fixed z,

$$P_{\widetilde{X}_k}\big\{\widetilde{X}_1(z) = 0\big\} \ge \prod_{i=1}^{s} P\big\{\widetilde{X}_1(z) = 0 | \widetilde{X}_0(w) = 1 \text{ in } \overline{B}(z, \rho) \text{ exactly}$$
$$\text{for the } w\text{'s in } A_i \cap \overline{B}(z, \rho)\big\}.$$

Finally, on the event $\{\widetilde{X}_k^{v_i}(w) = 1 \text{ for } w \in A_i\}$,

$$\begin{aligned} &P\big\{\widetilde{X}_1(z) = 0 | \widetilde{X}_0(w) = 1 \text{ in } \overline{B}(z, \rho) \text{ exactly for the } w\text{'s in } A_i \cap \overline{B}(z, \rho)\big\} \\ &\quad = P\big\{\widetilde{X}_{k+1}^{v_i}(z) = 0 | \widetilde{X}_k^{v_i}(w) = 1 \text{ in } \overline{B}(z, \rho) \text{ exactly} \\ &\qquad\qquad \text{for the } w\text{'s in } A_i \cap \overline{B}(z, \rho)\big\} \\ &\quad \ge P\{\widetilde{X}_{k+1}^{v_i}(z) = 0 | \widetilde{X}_k^{v_i}\} \quad [\text{by } (4.2)]. \end{aligned}$$

Thus, on the event $\{\widetilde{X}_k(w) = 1 \text{ on } C_k = \bigcup_1^s A_i\}$, which is contained in

$$\bigcap_{i=1}^{s} \{\widetilde{X}_k^{v_i}(w) = 1 \text{ for } w \in A_i\},$$

and under (4.21), it holds that

$$\begin{aligned} &P_x\big\{\widetilde{X}_{k+1}(z) = 0 \text{ for all } z \text{ in } A | \widetilde{X}_k, \widetilde{X}_k^v, v \in \mathbb{Z}^d\big\} \\ &\quad \ge \prod_{z \in A} \prod_{i=1}^{s} P\big\{\widetilde{X}_{k+1}^{v_i}(z) = 0 | \widetilde{X}_k^{v_i}\big\} \\ &\quad = \prod_{i=1}^{s} P\big\{\widetilde{X}_{k+1}^{v_i}(z) = 0 \text{ for all } z \in A | \widetilde{X}_k^{v_i}\big\} \\ &\quad \ge P_x\big\{Y_{k+1}(z) = 0 \text{ for all } z \in A | \widetilde{X}_k^v, v \in \mathbb{Z}^d\big\}. \end{aligned}$$

As before in the case $n = 1$, we can now couple $\widetilde{X}_{k+1}$ and $\widetilde{X}_{k+1}^v$ such that (4.21) continues to hold for $n = k + 1$. Equation (4.19) for $n = k + 1$ is immediate from (4.21) for $n \le k + 1$. □

PROOF OF PROPOSITION 1. We have to check (1.2), (1.4) and (1.22)–(1.24) for the chain $\{X_n\}$ defined in (1.36) [on the space S of (4.6)]. For the set $\mathscr{U}_2$ of (1.24) we shall take the singleton $\{\delta(\mathbf{0})\}$, that is, the state which has only one occupied site (which is automatically the leftmost particle and therefore shifted to the origin in X_n). For $\mathscr{U}_1$ in (1.23) we shall take the set

$$\mathscr{U}_1(K) := \{x \in S_0 = S \setminus \{\underline{0}\}: |x| \le K\} \tag{4.22}$$

for some large K, yet to be determined. Also, x_0 will be taken equal to $\delta(\mathbf{0})$.

Now first observe that from any state $x \in S$ the transition to x_0 in one step is possible. This (and several of the succeeding statements) is proven somewhat more simply in terms of the chain $\widetilde{X}_{\bullet}$ than in terms of $X_{\bullet}$ itself. Indeed, if $v_1, \dots, v_l$ are the occupied sites of a state $x \in \widetilde{S}$, then with z_0 as in (4.4),

$$\begin{aligned} P_x\{|\widetilde{X}_1| = 1\} &\ge P_x\{\widetilde{X}_1(z) = 1 \text{ only for } z = v_1 + z_0\} \\ &= P_x\{\widetilde{X}_1(v_1 + z_0) = 1\} \prod_{z \ne v_1 + z_0} P_x\{\widetilde{X}_1(z) = 0\}. \end{aligned}$$

Now for some constant $D_1 > 0$,

$$P_x\{\widetilde{X}_1(v_1 + z_0) = 1\} \ge D_1$$

for all x with $x(v_1) = 1$, by (4.2) and (4.4). Also $P_x\{\widetilde{X}_1(z) = 0\} = 1$ when $z \notin \overline{B}(v_i, \rho)$ for some $1 \le i \le l$ [by (1.35) and (4.4)]. For $z \in \overline{B}(v_i, \rho)$,

$$P_x\{\widetilde{X}_1(z) = 0\} = 1 - P_x\{\widetilde{X}_1(z) = 1\} \ge D_2 > 0$$

for some constant D_2 independent of x and z [by (1.33), (1.35) and (4.3)]. Thus

$$P_x\{|\widetilde{X}_1| = 1\} \ge D_1 D_2^{l(2\rho+1)^d}. \tag{4.23}$$

This proves the irreducibility condition (1.2). It also proves (1.24) (with $n_0 = 1$) for our choice of $\mathscr{U}_1$ and $\mathscr{U}_2$, because the estimate (4.23) depends on $l = |x|$ only; the right-hand side is at least

$$D_1 D_2^{K(2\rho+1)^d} \tag{4.24}$$

for any $x \in \mathscr{U}_1(K)$.

Condition (1.4), the certainty of absorption, follows directly from the assumption that $\widetilde{X}_{\bullet}$ is (strongly) subcritical and the fact that $\widetilde{X}_n(z) = 0$ automatically for all

$$z \notin \bigcup \overline{B}(v, n\rho),$$

where the union is over the finitely many v with $\widetilde{X}_0(v) \ne 0$ [cf. proof of (3.4)].

Condition (1.23) follows from Lemma 9. Indeed, if $x(v) = 1$ exactly when $v \in \{v_1, \dots, v_l\}$, with $l \le K$, then by (4.19),

$$\begin{aligned} P_x\{T > n\} &= P_x\{\widetilde{X}_n(z) = 1 \text{ for some } z\} \\ &\le P_x\{Y_n(z) = 1 \text{ for some } z\} \\ &\le \sum_{i=1}^{l} P_{\delta_{(v_i)}}\{\widetilde{X}_n(z) = 1 \text{ for some } z\} \\ &\le l P_{\delta(\mathbf{0})}\{T > n\} \le K P_{\delta(\mathbf{0})}\{T > n\}. \end{aligned}$$

Finally, condition (1.22) for any sufficiently large choice of K is guaranteed by Theorem 2. □

5. A central limit theorem. Finally we prove Theorem 3. We shall prove (1.41) for a single time only; that is, we take, $l = 1$ and $s_l = 1$. The general case follows then easily by induction on l, from the Markov property. Also, as before, we restrict ourselves to the aperiodic case (except when we prove $M = 0$, $\Sigma = \sigma^2 \times$ identity matrix for the discrete time contact process, which has period 2).

For the proof we shall make use of an analogue of the P^*-measure used in Lemma 8. However, this time we need a measure on the space $\widetilde{S}_0 := \widetilde{S} \setminus \{\underset{\sim}{0}\}$, which is the state space for $\widetilde{X}_\bullet$ minus its absorbing state. To define this we must first lift the function f of (1.17), which is defined on S_0, to $\widetilde{S}_0$. To this end let π be the projection from $\widetilde{S}_0$ to S_0. That is, if $\widetilde{x}(z) \neq 0$ exactly for $z \in \{\xi_1, \dots, \xi_\nu\}$, then we take $\pi(\widetilde{x}) = \widetilde{x} \oplus (-\xi_1)$, which has $\pi(\widetilde{x})(z) \neq 0$ if and only if $z \in \{\mathbf{0}, \xi_2 - \xi_1, \dots, \xi_\nu - \xi_1\}$. If $\widetilde{X}_n = \widetilde{x}$, then $\pi(\widetilde{x})$ is the corresponding state of X_n. We now define

$$\widetilde{f}(\widetilde{x}) = f(\pi(\widetilde{x})), \qquad \widetilde{x} \in \widetilde{S}_0.$$

It is easy to see that for $x = \pi(\widetilde{x})$,

$$\begin{aligned} &\sum_{\widetilde{y} \in \widetilde{S}} P(\widetilde{X}_n = \widetilde{y} \mid \widetilde{X}_0 = \widetilde{x}) \widetilde{f}(\widetilde{y}) \\ &\quad = E\{f(X_n) \mid \widetilde{X}_0 = \widetilde{x}\} = E\{f(X_n) \mid X_0 = x\} \\ &\quad = R^{-n} f(x) \quad \text{[see (1.17)]} \\ &\quad = R^{-n} \widetilde{f}(\widetilde{x}). \end{aligned}$$

Thus if we define for $C \in \widetilde{S}_0^{n+1}$,

$$\begin{aligned} &P^*\{\widetilde{X}_0, \widetilde{X}_1, \dots, \widetilde{X}_n \in C \mid \widetilde{X}_0 = \widetilde{x}\} \\ &\quad = \frac{R^n}{\widetilde{f}(\widetilde{x})} P\{\widetilde{X}_0, \dots, \widetilde{X}_n \in C \mid \widetilde{X}_0 = \widetilde{x}\} \widetilde{f}(\widetilde{X}_n), \end{aligned} \tag{5.1}$$

then P^* defines an honest distribution on the space of paths $\widetilde{S}_0^{\mathbb{Z}_+}$. (In order not to overburden notation we shall refrain from attaching asterisks to the $\widetilde{X}$

as well, in contrast to what we did in Lemma 8. Note also that the measure induced by P^* on the paths of the X-process agrees with the P^*-measure of Lemma 8.)

Now fix $\widetilde{x} \in \widetilde{S}_0$ and $y \in S_0$ for the remainder of the proof. Let $x = \pi(\widetilde{x}) \in S_0$ and let $B \subset \mathbb{Z}^d$. Then

$$\begin{aligned}
&P\{\xi_1(n) \in B,\ X_n = y | \widetilde{X}_0 = \widetilde{x}, T > n\}\\
&\quad = \frac{f(x)}{f(y)} \frac{1}{\widetilde{f}(\widetilde{x})} P\{\xi_1(n) \in B,\ X_n = y | \widetilde{X}_0 = \widetilde{x}\} f(y)[P_x\{T > n\}]^{-1}\\
&\quad \sim \Big[f(y) \sum_{w \in S_0} \mu(w)\Big]^{-1} \frac{R^n}{\widetilde{f}(\widetilde{x})} P\{\xi_1(n) \in B,\ X_n = y | \widetilde{X}_0 = \widetilde{x}\} f(y) \qquad (5.2)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{[by (1.26)]}\\
&\quad = \Big[f(y) \sum_{w \in S_0} \mu(w)\Big]^{-1} P^*\{\xi_1(n) \in B,\ \pi(\widetilde{X}_n) = y | \widetilde{X}_0 = \widetilde{x}\}.
\end{aligned}$$

We now define $\tau_0 = 0 < \tau_1 < \cdots$ as the successive times at which $X_\bullet$ visits x and write

$$W_i = (\widetilde{X}_{\tau_{i-1}}, \widetilde{X}_{\tau_{i-1}+1}, \dots, \widetilde{X}_{\tau_i}), \qquad i \geq 1, \tag{5.3}$$

for the ith excursion between visits to x. We already remarked in Lemma 8, that under P^*, $X_\bullet$ is positive-recurrent, so that a.e. $[P^*]$ all τ_i are finite and the W_i are well defined. Also by the strong Markov property, and the translation invariance property (1.33), the shifted excursions

$$\begin{aligned}
&W_i \oplus (-\xi_1(\tau_{i-1}))\\
&\quad := (\widetilde{X}_{\tau_{i-1}} \oplus (-\xi_1(\tau_{i-1})), \widetilde{X}_{\tau_{i-1}+1} \oplus (-\xi_1(\tau_{i-1})), \qquad (5.4)\\
&\qquad\qquad \dots, \widetilde{X}_{\tau_i} \oplus (-\xi_1(\tau_{i-1}))), \qquad i \geq 1,
\end{aligned}$$

are i.i.d. Define further

$$\Lambda_i = \text{“length” of } W_i = \tau_i - \tau_{i-1}, \tag{5.5}$$

$$\Delta_i = \text{“displacement” of } W_i = \xi_1(\tau_i) - \xi_1(\tau_{i-1}) \tag{5.6}$$

and

$$\Gamma_i = \text{“diameter” of } W_i = \max\{|\xi_1(n) - \xi_1(\tau_{i-1})|:\ \tau_{i-1} \leq n \leq \tau_i\}.$$

These are functions of W_i and, hence, $(\Lambda_i, \Delta_i, \Gamma_i)_{i\geq 1}$ are also i.i.d. Moreover (1.28) shows that the distribution of Λ_i under P^* has an exponentially bounded tail, as follows:

$$
\begin{aligned}
P^*\{\Lambda_i \geq l | \widetilde{X}_0 = \widetilde{x}\} &= P^*\{\Lambda_1 \geq l | \widetilde{X}_0 = \widetilde{x}\} \\
&= \sum_{k=l}^{\infty} P^*\{\Lambda_1 = k | \widetilde{X}_0 = \widetilde{x}\} = \sum_{k=l}^{\infty} P^*\{\tau_1 = k | \widetilde{X}_0 = \widetilde{x}\} \\
&= \sum_{k=l}^{\infty} R^k P\{X_k = x,\ X_r \neq x,\ 1 \leq r \leq k-1 | X_0 = x\} \\
&\leq \sum_{k=l}^{\infty} C_4 \left(\frac{R}{R+\eta(x)}\right)^k \leq D_1(x) \left(\frac{R}{R+\eta(x)}\right)^l.
\end{aligned} \tag{5.7}
$$

In addition we claim that for some $D_2 = D_2(x) < \infty$,

$$\Gamma_i \leq 2\Lambda_i \rho + D_2. \tag{5.8}$$

To see this, note that if $\widetilde{X}_n$ has occupied sites at $\xi_1(n), \dots, \xi_{\nu(n)}(n)$, then the occupied sites of $\widetilde{X}_{n+1}$ must all lie in

$$\bigcup_{i=1}^{\nu(n)} \overline{B}(\xi_i(n), \rho)$$

[by virtue of (1.35) and (1.31); compare (3.4)]. Therefore the occupied sites of $\widetilde{X}_n$ for any $\tau_{i-1} \leq n \leq \tau_i$ have to lie in

$$\bigcup_{i=1}^{\nu(\tau_{i-1})} \overline{B}(\xi_i(\tau_{i-1}), (\tau_i - \tau_{i-1})\rho)$$

and

$$
\begin{aligned}
\Gamma_i &\leq 2(\tau_i - \tau_{i-1})\rho + \max_{1 \leq r \leq \nu} \|\xi_r(\tau_{i-1}) - \xi_1(\tau_{i-1})\| \\
&\leq 2\Lambda_i \rho + \max_{r,s} \|v_r - v_s\|,
\end{aligned}
$$

where $v_1, v_2, \dots$ are the occupied sites of the state x (because $\widetilde{X}_{\tau_{i-1}} = \widetilde{x},\ X_{\tau_{i-1}} = x$) and $\nu = \nu(\tau_{i-1})$. Thus (5.8) holds with $D_2 = \max_{r,s} \|v_r - v_s\|$.

Next we define

$$\theta(n) = \max\{i\colon\ \tau_i \leq n\}.$$

Then

$$\xi_1(n) = \xi_1(\theta(n)) + \xi_1(n) - \xi_1(\theta(n)) = \xi_1(0) + \sum_{1}^{\theta(n)} \Delta_i + \xi_1(n) - \xi_1(\theta(n))$$

and

$$|\xi_1(n) - \xi_1(\theta(n))| \leq \Gamma(\theta(n)+1) \leq 2\Lambda_{\theta(n)+1}\rho + D_2.$$

In addition,

$$\frac{1}{\sqrt{r}}\Lambda_r \to 0 \quad \text{almost surely } [P^*]$$

because the Λ_r are i.i.d. and have all moments under P^*. Also, by the strong law of large numbers,

$$\frac{\theta(n)}{n} \to [E^*\{\Lambda_1 | \widetilde{X}_0 = \widetilde{x}\}]^{-1} \quad \text{almost surely } [P^*], \tag{5.9}$$

again because the Λ_r are i.i.d. and $\tau_i = \sum_1^i \Lambda_r$. Thus

$$\frac{\Lambda_{\theta(n)+1}}{\sqrt{n}} \to 0 \quad \text{almost surely } [P^*]. \tag{5.10}$$

The standard proof of the central limit theorem for Markov chains [Chung (1967), Section I.16] now shows that

$$\frac{\xi_1(n) - nM}{\sqrt{n}} \Rightarrow N(\mathbf{0}, \Sigma) \tag{5.11}$$

under P^*, conditioned on $X_0 = \widetilde{x}$, with

$$M = \frac{E^*\{\Delta_1 | \widetilde{X}_0 = \widetilde{x}\}}{E^*\{\Lambda_1 | \widetilde{X}_0 = \widetilde{x}\}}, \tag{5.12}$$

$$\Sigma(i, j) = \frac{1}{E^*\{\Lambda_1 | \widetilde{X}_0 = \widetilde{x}\}} E^*\{(\Delta_{1,i} - M_i\Lambda_1)(\Delta_{1,j} - M_j\Lambda_1) | \widetilde{X}_0 = \widetilde{x}\}, \tag{5.13}$$

where $\Delta_{1,i}$ and M_i are the ith component of Δ_1 and M, respectively. Slightly more explicit expressions for M and Σ can be given in the same form as in Theorems I.14.5, I.14.7 and its Corollary in Chung (1967). Note in this connection that

$$\begin{aligned} &E^*\{\text{number of indices } 0 \le n < \tau_1 \text{ with } X_n = v | \widetilde{X}_0 = \widetilde{x}\} \\ &\quad = E^*\{\text{number of indices } 0 \le n < \tau_1 \text{ with } X_n = v | X_0 = x\} \\ &\quad = \frac{f(v)\mu(v)}{f(x)\mu(x)} \end{aligned} \tag{5.14}$$

because $f(v)\mu(v)$ is the unique invariant probability measure for the $X_\bullet$ chain under P^*. That is,

$$\begin{aligned} &\sum_{x \in S_0} f(x)\mu(x) P^*\{X_1 = v | X_0 = x\} \\ &\quad = Rf(v) \sum_{x \in S_0} \mu(x) P\{X_1 = v | X_0 = x\} = f(v)\mu(v) \end{aligned} \tag{5.15}$$

[by (1.18)] and (1.19) holds. We shall not pursue more explicit expressions for M and Σ here; see, however, the end of this proof.

In view of (5.11) we basically only have to prove the independence of X_n and $[\xi_1(n) - nM]/\sqrt{n}$. The arguments for this are at least part of the folklore and we shall therefore be brief. For $\varepsilon > 0$ we can choose n_5 such that for all $n \geq 2n_5$,

$$\begin{aligned} &|P^*\{X_n = y | \widetilde{X}_0 = \widetilde{x}\} - f(y)\mu(y)| \\ &\qquad = |P^*\{X_n = y | X_0 = x\} - f(y)\mu(y)| \leq \varepsilon \end{aligned} \tag{5.16}$$

and

$$P^*\{\tau(\theta(n - 2n_5) + 1) \geq n - n_5 | \widetilde{X}_0 = \widetilde{x}\} \leq \varepsilon. \tag{5.17}$$

Equation (5.16) is immediate from (1.26) and (1.27) after a translation to the P^*-measure. In (5.17) we have written $\tau(\theta(n - 2n_5) + 1)$ instead of $\tau_{\theta(n-2n_5)+1}$ for typographical convenience; the estimate (5.17) itself is straightforward renewal theory, since $\tau(\theta(n-2n_5)+1)$ is the smallest $m > n-2n_5$ with $X_m = x$ [see Chung (1967), Theorem I.14.2]. It is further easy to see by means of (5.8) and Chung [(1967), Theorem I.14.2] that for fixed n_5,

$$\lim_{a \to \infty} P^*\{|\xi_1(n) - \xi_1(\tau(\theta(n - 2n_5) + 1))| > a | \widetilde{X}_0 = \widetilde{x}\} = 0$$

uniformly in $n \geq 2n_5$. Therefore,

$$\begin{aligned} &\frac{1}{\sqrt{n}} |\xi_1(n) - \xi_1(\tau(\theta(n - 2n_5) + 1))| \to 0 \quad \text{in } P^*\text{-measure,} \\ &\qquad\qquad\qquad \text{conditioned on } \widetilde{X}_1(0) = \widetilde{x}, \text{ as } n \to \infty. \end{aligned} \tag{5.18}$$

Finally, a decomposition with respect to the value of $\tau(\theta(n - 2n_5) + 1)$ gives

$$\begin{aligned} &|P^*\{\xi_1(\tau(\theta(n - 2n_5) + 1)) \in B,\ X_n = y | \widetilde{X}_0 = \widetilde{x}\} \\ &\qquad - P^*\{\xi_1(\tau(\theta(n - 2n_5) + 1)) \in B | \widetilde{X}_0 = \widetilde{x}\} f(y)\mu(y)| \\ &\quad \leq P^*\{\tau(\theta(n - 2n_5) + 1) \geq n - n_5 | \widetilde{X}_0 = \widetilde{x}\} \\ &\qquad + \sum_{n-2n_5 < r < n-n_5} P^*\{\tau(\theta(n - 2n_5) + 1) = r,\ \xi_1(r) \in B | \widetilde{X}_0 = \widetilde{x}\} \\ &\qquad\qquad \times |P^*\{X_{n-r} = y | X_0 = x\} - f(y)\mu(y)| \\ &\quad \leq 2\varepsilon \quad [\text{by (5.16) and (5.17)}]. \end{aligned} \tag{5.19}$$

Since $\varepsilon > 0$ is arbitrary, it is now not hard to obtain from (5.11), (5.18) and (5.19) that

$$\begin{aligned} &\lim_{n \to \infty} P^* \left\{ \frac{\xi_1(n) - nM}{\sqrt{n}} < \gamma, X_n = y | \widetilde{X}_0 = \widetilde{x} \right\} \\ &\qquad = P\{G_1 \leq \gamma\} f(y)\mu(y). \end{aligned}$$

Translating this back to the P-measure gives (1.41) for $l = 1,\ s_l = 1$ [by means of (1.26) again].

The last statement of Theorem 3 is about the values of M and Σ for the discrete time contact process. Since M and Σ are independent of the choice of $\tilde{x}$ [see Chung (1967), Corollary to Theorem I.15.4 and Corollary 2 to Theorem I.16.1], we can take $\tilde{x} = \delta(\mathbf{0})$. It is then clear from the symmetry of the discrete time contact process that $M = \mathbf{0}$ and $\Sigma_{ij} = 0$ for $i \neq j$ [see (5.12) and (5.13)]. Also all $\Sigma_{i,i}$, $1 \leq i \leq d$, must have the same value, say σ^2, so the only nontrivial part of our claim is that

$$\sigma^2 = \frac{1}{E^*\{\Lambda_1 | \widetilde{X}_0 = \delta(\mathbf{0})\}} E^*\{(\Delta_{1,1} - M_1\Lambda_1)^2 | \widetilde{X}_0 = \delta(\mathbf{0})\} > 0$$

or, equivalently, that

$$P^*\{\Delta_{1,1} - M_1\Lambda_1 \neq 0 | \widetilde{X}_0 = \delta(\mathbf{0})\} > 0.$$

However, $M_1 = 0$ and one easily sees that

$$\begin{aligned} &P^*\{\Delta_{1,1} \neq 0 | \widetilde{X}_0 = \delta(\mathbf{0})\} \\ &\quad \geq P^*\{\tau_1 = 2, \Delta_{1,1} = 2 | \widetilde{X}_0 = \delta(\mathbf{0})\} \\ &\quad \geq P^*\{\widetilde{X}_1 = \delta(e_1) + \delta(-e_1),\ \widetilde{X}_2 = \delta(2e_1) | \widetilde{X}_0 = \delta(\mathbf{0})\} > 0, \end{aligned}$$

where $e_1 = (1, 0, \ldots, 0)$. Indeed the event $\{\widetilde{X}(1) = \delta(e_1) + \delta(-e_1)\}$ occurs if at time 1 exactly the sites e_1 and $-e_1$ are occupied; similarly $\{\widetilde{X}(2) = \delta(2e_1)\}$ is the event that only the site $2e_1$ is occupied at time 2. The description in Section 4 of the discrete time contact process shows that these events have strictly positive probability. Thus $\sigma^2 > 0$ and the proof of Theorem 3 is complete.

Acknowledgments. P. F. and H. K. thank the Newton Institute of Mathematical Sciences for its hospitality and its support during the fall of 1993, when this research was carried out.

REFERENCES

AIZENMAN, M. and BARSKY, D. J. (1987). Sharpness of the phase transition in percolation models. *Comm. Math. Phys.* **108** 489–526.

BEZUIDENHOUT, C. and GRIMMETT, G. (1991). Exponential decay for subcritical contact and percolation processes. *Ann. Probab.* **19** 984–1009.

CAMPANINO, M., CHAYES, J. T. and CHAYES, L. (1991). Gaussian fluctuations of connectivities in the subcritical regime of percolation. *Probab. Theory Related Fields* **88** 269–341.

CHUNG, K. L. (1967). *Markov Chains with Stationary Transition Probabilities*, 2nd ed. Springer, New York.

DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth and Brooks/Cole, Belmont, CA.

FERRARI, P. A., KESTEN, H., MARTÍNEZ, S. and PICCO, P. (1995). Existence of quasi–stationary distributions. A renewal dynamical approach. *Ann. Probab.* **23** 501–521.

GRIMMETT, G. (1989). *Percolation*. Springer, New York.

KESTEN, H. (1995). A ratio limit theorem for (sub) Markov chains on $\{1, 2, \ldots\}$ with bounded jumps. *Adv. in Appl. Probab.* **27** 652–691.

NGUYEN, B. G. and YANG, W-S. (1995). Gaussian limit for critical oriented percolation in high dimensions. *J. Statist. Phys.* **78** 841–876.

PAKES, A. (1995). Quasi-stationary laws for Markov processes: examples of an always proximate absorbing state. *Adv. in Appl. Probab.* **27** 120–145.

ROBERTS, G. O. and JACKA, S. D. (1994). Weak convergence of conditioned birth and death processes. *J. Appl. Probab.* **31** 90–100.

SENETA, E. and VERE-JONES, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3** 403–434.

VERE-JONES, D. (1967). Ergodic properties of nonnegative matrices I. *Pacific J. Math.* **22** 361–396.

YAGLOM, A. M. (1947). Certain limit theorems of the theory of branching stochastic processes. *Dokl. Akad. Nauk SSSR* **56** 797–798.

P. A. FERRARI
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE DE SÃO PAULO
CX. POSTAL 20570
01498-970, SÃO PAULO SP
BRASIL
E-MAIL: pablo@ime.nsp.br

H. KESTEN
DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853-7901
E-MAIL: kesten@math.cornell.edu

S. MARTÍNEZ
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA
UNIVERSIDAD DE CHILE
CASILLA 170
CORREO 3
SANTIAGO 3
CHILE
E-MAIL: smartine@dim.uchile.cl