ON SOME GROWTH MODELS WITH A SMALL PARAMETER

Harry Kesten and Roberto H. Schonmann*

Abstract. We consider the behavior of the asymptotic speed of growth and the asymptotic shape in some growth models, when a certain parameter becomes small. The basic example treated is the variant of Richardson's growth model on $\mathbb{Z}^d$ in which each site which is not yet occupied becomes occupied at rate 1 if it has at least two occupied neighbors, at rate $\varepsilon \leq 1$ if it has exactly 1 occupied neighbor and, of course, at rate 0 if it has no occupied neighbor. Occupied sites remain occupied forever. Starting from a single occupied site, this model has asymptotic speeds of growth in each direction and these speeds determine an asymptotic shape in the usual sense. It is proven that as ε tends to 0, the asymptotic speeds scale as $\varepsilon^{1/d}$ and the asymptotic shape, when renormalized by dividing it by $\varepsilon^{1/d}$, converges to a cube. Other similar models which are partially oriented are also studied.

1. Introduction. In this paper we consider variants of some typical growth models, in which now the rule depends on some parameter, and we ask how quantities of interest as the asymptotic speed of growth and the asymptotic shape behave in case this parameter becomes very small. To be more specific we describe now the basic example that we have in mind. This is the growth model on the lattice $\mathbb{Z}^d$, in which a vacant site becomes occupied at exponential rate 1 if it has at least 2 occupied neighbors and at rate $\varepsilon \leq 1$ if it has exactly 1 occupied neighbor. For this model and all other ones considered in this paper, we suppose, as usual, that occupied sites are never vacated, that initially only the origin of $\mathbb{Z}^d$ is occupied and that no other site can become occupied before any of its neighbors becomes occupied (sites which have no occupied neighbor are becoming occupied at rate 0). Throughout we take $\varepsilon \leq 1$ and we shall actually focus on the regime where ε is very small. We will refer to this model as the *basic model*.

As motivation for such a model one can think of a crystal growing in a supersaturated solution. New atoms are added to the seed by mechanisms which certainly should make the probability of the absorption of a new atom at a given position dependent on the number of sites adjacent to this position which already contain an atom of the crystal. It is not unreasonable to expect that the creation of protuberances (the absorption of an atom at a lattice position which has only one

---

1991 *Mathematics Subject Classification.* Primary 60K35.

Key words and phrases. Growth models, asymptotic speed, asymptotic shape, Eden model, Richardson model, first-passage percolation.

*The work of R.H.S was supported by the N.S.F through grant DMS 91-00725. In addition, both authors were supported by the Newton Institute in Cambridge. The authors thank the Newton Institute for its support and hospitality.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

neighboring position already belonging to the crystal) is a mechanism which is far slower than the absorption of further atoms at positions that already have at least 2 occupied lattice neighbors. Of course, one can also modify our simple model and suppose that the rates of absorption depend in a more elaborate fashion on the number of occupied neighbors, but for simplicity we will restrict ourselves to the model defined above and some partially oriented variants.

Still, as motivation, it is worth pointing out that the problems treated in this paper may be also useful in the analysis of the relaxation patterns of stochastic Ising models under external magnetic field, at very low temperatures, as for instance discussed in Section 8 of Schonmann (1984). In that paper relaxation mechanisms for such systems in such a regime are reviewed, and in particular it is stressed that if one is considering infinite or very large finite systems, then not only the rate of appearance of droplets of the stable phase in the background given by the metastable phase, but also the speed with which these droplets grow, is crucial to analyze the global relaxation patterns. The results derived in the present paper refer to systems which are far simpler than the stochastic Ising models, but the results obtained and the methods developed here may be the first step towards sharpening certain results on the droplet driven relaxation of stochastic Ising models. For instance, they suggest that in Theorem 6 in Schonmann (1984) the constant C_2 is sharp (rather than C_1).

Our model is also closely related to Eden's and Richardson's models (see Eden (1961) and Richardson (1973)). In Eden's model one studies a sequence of connected sets B_n of $\mathbb{Z}^d$. B_n consists of n vertices and contains the origin ($B_1 = \mathbf{0}$). B_{n+1} is formed from B_n by adding a single vertex y_n from

$$\partial B_n := \{x : x \notin B_n, \text{ but } x \text{ is adjacent to } B_n\}.$$

In one of the common versions of this model one chooses y_n uniformly from ∂B_n. It is known (see Richardson (1973)) that the sequence of sets $B_1, B_2, \dots$ has the same distribution as the sequence of sets $\widetilde{B}(t_1), \widetilde{B}(t_2), \dots$ obtained in the following Richardson model. At time 0 only the origin is occupied. A vertex $x \neq \mathbf{0}$ cannot become occupied before one of its neighbors is occupied. When x has at least one occupied neighbor, then x itself becomes occupied at rate 1. $\widetilde{B}(t)$ denotes the set of vertices occupied at time t, and t_n is the first time when exactly n vertices are occupied. Other versions of the Eden model pick y_n from ∂B_n with a probability which is proportional to the number of neighbors of y_n in ∂B_n, that is

$$\mathbb{P}\{y_n = x | B_1, B_2, \dots, B_n\} = \frac{1}{Z_n}(\text{number of neigbors of } x \text{ in } \partial B_n),$$

where

$$Z_n = \sum_{x \in \partial B_n} (\text{number of neigbors of } x \text{ in } \partial B_n).$$

For the corresponding Richardson model, x should become occupied at a rate proportional to the number of its neighbors which are already occupied. Our basic model corresponds to an Eden model in which y_n is taken equal to a vertex $x \in \partial B_n$

with probability $\varepsilon/\widetilde{Z}_n$ if x has only one neighbor in B_n and with probability $1/\widetilde{Z}_n$ if x has at least two neighbors in B_n, where now

$$\widetilde{Z}_n = \widetilde{Z}_n(\varepsilon) = \sum_{x \in \partial B_n} [\varepsilon + (1-\varepsilon)\mathbb{1}(x \text{ has at least 2 neighbors in } B_n)].$$

(Here and in the sequel $\mathbb{1}(A)$, or sometimes $\mathbb{1}_A$, denotes the indicator function of A.) The corresponding version of Richardson's model is described in detail at the beginning of the next section. In Remark 2.1 we also mention briefly a generalization of the basic model in the context of first–passage percolation.

We review now some basic facts which are known to hold for the growth model we are considering. For proofs and more details see the beginning of Section 2 below and also Section 11c of Durrett (1988) and references therein and Schürger (1979). We will denote by η_t the indicator function of the set of sites occupied at time t. The asymptotic speed of growth can be defined as a function of the direction, but for the moment we will consider only the speed in a coordinate direction. Let $\mathbf{e}_i$ be the i-th coordinate vector $(0, 0, \ldots, 1, \ldots, 0)$, (with the 1 in the i–th place) and define the random times

$$T(x) := \inf\{t > 0 : \eta_t(x) = 1\}. \tag{1.1}$$

Then almost surely

$$\frac{n}{T(n\mathbf{e}_1)} \to v(\varepsilon, d), \tag{1.2}$$

as $n \to \infty$, where $v(\varepsilon, d)$ is a constant which depends on ε and will be called *the speed of growth.*

In order to introduce the notion of asymptotic shape, one first defines the fattened occupied region

$$\overline{\eta}_t := \{z \in \mathbb{R} : ||z - x||_\infty \le 1/2 \text{ for some } x \text{ with } \eta_t(x) = 1\},$$

where $||y||_\infty = \max_{1 \le i \le d} |y_i|$ when $y = (y_1, \ldots, y_d)$.

For $U \subset R^d$ and $\alpha \in R$, we write $\alpha U := \{\alpha x : x \in U\}$. The following is then true. There exists a compact convex set $A(\varepsilon) \in \mathbb{R}$ such that for every $\delta > 0$,

$$\mathbb{P}\{(1-\delta)tA(\varepsilon) \subset \overline{\eta}_t \subset (1+\delta)tA(\varepsilon) \text{ for all large } t\} = 1. \tag{1.3}$$

Two trivial bounds can be found for $v(\varepsilon, d)$. By comparison with a model in which only sites on the x-axis can be occupied, or the case in which $\varepsilon = 1$, one obtains, respectively, the following lower and upper bounds.

$$C_1\varepsilon \le v(\varepsilon, d) \le C_2,$$

where C_1 and C_2 are strictly positive finite constants. It is natural to ask what the actual scaling of $v(\varepsilon, d)$ is when $\varepsilon \downarrow 0$. The answer is neither one of the trivial bounds above.

Theorem 1.1. *For the basic model in dimension d,*

$$C_1\varepsilon^{1/d} \le v(\varepsilon, d) \le C_2\varepsilon^{1/d},$$

where C_1 and C_2 are strictly positive finite constants.

As for the asymptotic shape, we will show that, when properly rescaled, it is close to a cube when ε is small. The notion of 'closeness' here is the same one that appeared in (1.3) above. Unfortunately, we are only able to show a result which is somewhat weaker than convergence to a cube, but which is strong enough to assure the existence of convergence along suitable subsequences of ε's which go to 0. The trouble is that we were not able to rule out the possibility that for different such sequences the sidelength of the limiting cube is different. This leads to the following explicit conjecture.

Conjecture 1. *In the basic model*

$$\lim_{\varepsilon\downarrow 0} \varepsilon^{-1/d} v(\varepsilon, d) \in (0, \infty)$$

exists.

To state the precise shape result we define

$$Q(\ell) := [-\ell, \ell]^d. \tag{1.4}$$

Theorem 1.2. *For the basic model in dimension d, there exist strictly positive finite constants C_1 and C_2 and a function $\ell : [0,\infty) \to [C_1, C_2]$ such that for every $\delta > 0$ there exists an ε_0, such that for all $\varepsilon < \varepsilon_0$,*

$$(1-\delta)Q(\ell(\varepsilon)) \;\subset\; \varepsilon^{-1/d}A(\varepsilon) \;\subset\; (1+\delta)Q(\ell(\varepsilon)).$$

The following is an immediate consequence.

Corollary 1.1. *For the basic model in dimension d, for every sequence $\varepsilon_i \downarrow 0$ there exists a subsequence ε_{i_k} and a number $0 < \ell < \infty$, such that $(\varepsilon_{i_k})^{-1/d}A(\varepsilon_{i_k}) \to Q(\ell)$, in the sense that for all $\delta > 0$ there exists a $k_0 < \infty$ for which*

$$(1-\delta)Q(\ell) \;\subset\; (\varepsilon_{i_k})^{-1/d}A(\varepsilon_{i_k}) \;\subset\; (1+\delta)Q(\ell), \quad k \ge k_0.$$

Another consequence of Theorem 1.2 is that the asymptotic speed in any direction scales also as $\varepsilon^{1/d}$. Moreover, at any time, the furthest occupied point in the first coordinate direction is not much further than the furthest occupied point on the first coordinate axis itself. The first fact can also be proven more directly, using the arguments in the proof of Theorem 1.1.

Remark 1.1. R. Kotecký asked us what happens if the vertex x becomes occupied at rate $\varepsilon(j)$ if it has a single occupied neighbor in the j–th coordinate direction (positive or negative), and at rate 1 if it has two or more occupied neighbors (in any direction). Thus occupation spreads at different rates in different directions in this version. Abbreviate $(\varepsilon(1), \dots, \varepsilon(d))$ to $\overline{\varepsilon}$. Our proofs then carry over essentially unchanged to show that there exists an asymptotic shape $A(\overline{\varepsilon})$ (as $t \to \infty$), and the following analogue of Theorem 1.2 holds:

There exist strictly positive finite constants C_1 and C_2 and functions $\ell_j : [0,\infty)^d \to [C_1, C_2]$ such that for every $\delta > 0$ there exists an ε_0, such that for all $\varepsilon(j) < \varepsilon_0$,

$$(1-\delta)\prod_{j=1}^{d}[-\ell_j(\overline{\varepsilon})\varepsilon(j)^{1/d},\ \ell_j(\overline{\varepsilon})\varepsilon(j)^{1/d}] \subset A(\overline{\varepsilon})$$
$$\subset (1+\delta)\prod_{j=1}^{d}[-\ell_j(\overline{\varepsilon})\varepsilon(j)^{1/d},\ \ell_j(\overline{\varepsilon})\varepsilon(j)^{1/d}].$$

In particular, for very different $\varepsilon(j)$ the asymptotic shape will be near to some kind of rectangular box, rather than to a cube.

Remark 1.2. It is worth stressing that it is not hard for one to become convinced that Theorem 1.2 should be true by the wrong reasoning! In the first stages of the growth of the system, the occupied region is likely to be most of the time a rectangle, because for a first protuberance to be created at the surface of a small rectangle the time needed is of order ε^{-1}, while the time needed for a new rectangle to be reached after such a protuberance appears is of order 1. (By 'rectangle' we mean a product of lattice intervals.) If the rectangle is far from being a cube, then it is more likely that the next protuberance will appear adjacent to one of the larger faces, and will help the growing occupied region to approach a cubic shape. While the statements above are correct, they only refer to times which are short compared to ε^{-1}. Our interest, on the other hand, is in the behavior of the system with fixed small ε after very long times. One way to say it is that we are interested in first letting $t \to \infty$ and then letting $\varepsilon \downarrow 0$. When the linear size of the growing occupied region becomes comparable to $\varepsilon^{-1/(d-1)}$, protuberances appear throughout the surface at a rate of order 1, and the simple picture described above breaks down. In fact when the occupied region is roughly a cube of edge size L, protuberances will appear on its surface at a rate εL^{d-1}.

Theorem 1.2 gives some information about the asymptotic shape in the particular case in which ε is small. It is worth pointing out that few results are available about the asymptotic shapes of growth models. A well known exception is the 'flat edge' result of Durrett and Liggett (1981).

We next describe our results for a (partially) oriented generalization of the basic model. Let $\mathcal{N}$ be some finite subset of $\mathbb{Z}^d$ which does not contain the origin. The basic model is then of the following form: an unoccupied vertex x becomes occupied at rate 0 if there are no occupied vertices in $x+\mathcal{N}$, at rate ε if there is exactly one occupied vertex in $x+\mathcal{N}$, and at rate 1 if there are at least two occupied vertices in $x+\mathcal{N}$. The set $\mathcal{N}$ for the basic model is

$$\mathcal{N}_0 := \{\pm\mathbf{e}_i : 1 \le i \le d\}.$$

We now replace this by

$$\mathcal{N}_\nu := \{\pm\mathbf{e}_i : 1 \le i \le d-\nu\} \cup \{\mathbf{e}_j : d-\nu+1 \le j \le d\}.$$

Thus, in this case occupation of vertices can 'spread only in the positive direction' for the last ν coordinate directions. $T(x)$ is still defined by (1.1) and the following limits still exists almost surely (compare (1.2)):

$$v(\varepsilon, d, \nu) = \lim_{n \to \infty} \frac{n}{T_{n\mathbf{e}_j}}, \quad 1 \le j \le d - \nu,$$

and

$$w(\varepsilon, d, \nu) = \lim_{n \to \infty} \frac{n}{T_{n\mathbf{e}_j}}, \quad d - \nu + 1 \le j \le d.$$

These limits do not depend on j in the indicated ranges because of obvious symmetry properties of the model. $v(\varepsilon, d, \nu)$ and $w(\varepsilon, d, \nu)$ are the speeds at which occupation spreads along the coordinate axes. In analogy with concepts in first–passage percolation we will call these *point–to–point* speeds, since they measure the time needed for the occupation to spread from the point $\mathbf{0}$ to the point $n\mathbf{e}_j$. Also of interest are *point–to–hyperplane* speeds which should be defined as follows. In the above model let

$$\begin{aligned}\sigma(n, j, \varepsilon, \nu) &= \inf\{T(x) : x_j = n\} \\ &= \text{first time a point in the hyperplane } \{x : x_j = n\} \text{ becomes occupied.}\end{aligned}$$

Then the point–to–hyperplane speeds should be

$$s(\varepsilon, \nu) = \lim_{n \to \infty} \frac{n}{\sigma(n, j, \varepsilon, \nu)} \quad \text{for } 1 \le j \le d - \nu,$$

$$t(\varepsilon, \nu) = \lim_{n \to \infty} \frac{n}{\sigma(n, j, \varepsilon, \nu)} \quad \text{for } d - \nu + 1 \le j \le d.$$

Unfortunately we have at the moment no complete proof that these limits exist when $\nu > 0$. We therefore only give results about

$$\limsup_{n \to \infty} \frac{n}{\sigma(n, j, \varepsilon, \nu)} \quad \text{and} \quad \liminf_{n \to \infty} \frac{n}{\sigma(n, j, \varepsilon, \nu)}$$

in the next theorem.

In the basic model with all coordinates unoriented ($\nu = 0$) it follows from the argument of Cox and Durrett (1981), proof of Theorem 5, that the limit for $s(\varepsilon, 0)$ exists and that

$$s(\varepsilon, 0) = v(\varepsilon, d, 0).$$

When exactly one coordinate is oriented ($\nu = 1$), then the same argument can be used to show that the limit for $t(\varepsilon, 1)$ exists and

$$t(\varepsilon, 1) = w(\varepsilon, d, 1),$$

i.e., the point–to–point and point–to–hyperplane speeds are the same for the single oriented direction. However, for general ν neither $s(\varepsilon, \nu) = v(\varepsilon, d, \nu)$ nor $t(\varepsilon, \nu) = w(\varepsilon, d, \nu)$ are expected to hold and our next theorem implies that the former definitely fails for small ε.

Theorem 1.3. *For* $\nu < d$

$$v(\varepsilon, d, \nu) = v(\varepsilon, d-\nu, 0) \tag{1.5}$$

and there exist constants $0 < C_1, C_2 < \infty$ *such that*

$$C_1\varepsilon^{1/(d-\nu)} \leq v(\varepsilon, d, \nu) \leq C_2\varepsilon^{1/(d-\nu)}. \tag{1.6}$$

Moreover for $0 < \nu \leq d$

$$w(\varepsilon, d, \nu) = w(\varepsilon, d-\nu+1, 1) \tag{1.7}$$

and there exist constants $0 < C_1, C_2 < \infty$ *such that*

$$C_1\varepsilon^{1/(d-\nu+1)} \leq w(\varepsilon, d, \nu) \leq C_2\varepsilon^{1/(d-\nu+1)}. \tag{1.8}$$

Finally

$$\begin{aligned} C_1\varepsilon^{\frac{d+1}{d(d-\nu+1)}} &\leq \liminf_{n\to\infty} \frac{n}{\sigma(n,j,\varepsilon,\nu)} \\ &\leq \limsup_{n\to\infty} \frac{n}{\sigma(n,j,\varepsilon,\nu)} \leq C_2\varepsilon^{\frac{d+1}{d(d-\nu+1)}} \ \textit{for}\ 1 \leq j \leq d-\nu, \end{aligned} \tag{1.9}$$

and

$$\begin{aligned} C_1\varepsilon^{1/(d-\nu+1)} &\leq \liminf_{n\to\infty} \frac{n}{\sigma(n,j,\varepsilon,\nu)} \\ &\leq \limsup_{n\to\infty} \frac{n}{\sigma(n,j,\varepsilon,\nu)} \leq C_2\varepsilon^{1/(d-\nu+1)} \ \textit{for}\ d-\nu+1 \leq j \leq d. \end{aligned} \tag{1.10}$$

Remark 1.3. (1.9) and (1.10) show that the process moves faster (in the sense of point–to–hyperplane speeds) in the oriented than in the unoriented directions. We have no intuitive argument for this fact. Note also that the point–to–hyperplane speed in the unoriented directions (given by (1.9)) is larger than the point–to–point speed in the totally unoriented model in dimension $d-\nu$, because

$$\frac{d+1}{d(d-\nu+1)} < \frac{1}{d-\nu}.$$

Thus, the oriented directions still 'help' the process to move in the unoriented directions.

Further variants of the basic model can be obtained by changing the rates at which points are occupied. For instance we considered the totally unoriented model in which a vertex x becomes occupied at rate ε when one or two neighbors of x are occupied and at rate 1 when at least three neighbors of x are occupied. We were unable to prove the following conjecture.

Conjecture. *For the above rates there exist constants* $0 < C_1 \leq C_2 < \infty$ *such that*

$$C_1 \varepsilon \leq \lim_{n \to \infty} \frac{n}{T_{n\mathbf{e}_1}} \leq C_2 \varepsilon. \tag{1.11}$$

The lefthand inequality is trivial; only the right hand inequality is significant. The conjecture is true for $d = 2$, but for $d = 3$ we could only prove the following weaker result.

Theorem 1.4. *If* $d = 3$ *and if* x *becomes occupied at rate* ε *when it has* $1, 2$ *or* 3 *occupied neighbors, and at rate* 1 *when it has at least* 4 *occupied neighbors, then (1.11) holds.*

Since this seems a far from optimal result we shall not prove this theorem here. We merely state here that the proof rests on showing that if $x = (x_1, \dots, x_d)$ is occupied at a certain time t, then at t there exists a chronological path from $\mathbf{0}$ to x of ℓ vertices which contains at least $C_3 \ell$ slow points for some ℓ (see the lines before Lemma 2.3 for the terminology). From this one concludes that x will not become occupied before time $C_4 \ell / \varepsilon$ outside an event of negligeable probability. Since necessarily $\ell \geq \max_j |x_j|$, we conclude that w. p. 1, $T(x)$ is at least $C_4 ||x||_\infty / \varepsilon = C_4 \max_j |x_j| / \varepsilon$ for all large x. The rules for constructing the chronological path are, however, much more complicated in this case than in the proof of Lemma 2.3

As usual we will use C_0, $C_1 \dots$ to denote strictly positive finite constants whose precise value is irrelevant and may change from appearance to appearance. All of these will be *independent of* ε, but they may depend on the dimension d. Some further notation which we shall use is

$$\mathbb{1}(A) \text{ or } \mathbb{1}_A = \text{indicator function of } A;$$

$$||x||_1 = \sum_{i=1}^{d} |x_i| \text{ and } ||x||_\infty = \max_{1 \leq i \leq d} |x_i| \text{ when } x = (x_1, \dots, x_d);$$

$$a \wedge b = \min(a, b);$$

$$\mathbf{0} = \text{the origin and } \mathbf{e}_i = \text{i–th coordinate vector.}$$

2. Proof of Theorem 1.1.

It is convenient to construct our basic growth model by means of exponential clocks. To each site $x \in \mathbb{Z}^d$ associate two independent exponential random variables, $\tau^{(1)}(x)$ and $\tau^{(2)}(x)$, with respective means $1/\varepsilon$ and $1/(1 - \varepsilon)$. We suppose that all these random variables associated to the sites of $\mathbb{Z}^d$ are independent. Recall the definition (1.1), and let $V^{(i)}(x)$, $i = 1, 2$, be the first time when the site x has at least i occupied neighbors. Then the process which starts with only the origin occupied can be constructed recursively by setting $T_{\mathbf{0}} = 0$ and for $x \neq \mathbf{0}$,

$$T(x) = \min\{V^{(1)}(x) + \tau^{(1)}(x), V^{(2)}(x) + \tau^{(2)}(x)\}. \tag{2.1}$$

More precisely, let $B(0) = \{\mathbf{0}\}$ and

$$\begin{aligned} B(t) &= \{x : x \text{ is occupied at time } t\} \\ &= \{x : T(x) \leq t\}, \end{aligned}$$

$$\partial B(t) = \{x : x \text{ is adjacent to } B(t) \text{ but } x \notin B(t)\}.$$

Each $x \in \partial B(t)$ has at least one occupied neighbor in $B(t)$, so that $V^{(1)}(x) \le t$. If x has two or more neighbors in $B(t)$, then even $V^{(2)}(x) \le t$. Let

$$\partial_1 = \{x \in \partial B(t) : x \text{ has exactly one neighbor in } B(t)\},$$

$$\partial_2 = \{x \in \partial B(t) : x \text{ has at least two neighbors in } B(t)\}.$$

Now for a given t and $B(t)$ take $B(s) = B(t)$ for all

$$s < S = S(t) := \min\{V^{(1)}(x) + \tau^{(1)}(x) : x \in \partial_1\} \wedge \min\{V^{(2)}(x) + \tau^{(2)}(x) : x \in \partial_2\}.$$

At time $S, B(t)$ changes to

$$B(S) = B(t) \cup \{x\},$$

where $x \in \partial B(t)$ is such that $S = V^{(1)}(x) + \tau^{(1)}(x)$ or $S = V^{(2)}(x) + \tau^{(2)}(x)$. $S = T(x)$ for this x. Due to this change from $B(t)$ to $B(S)$, also ∂B changes and so do ∂_1 and ∂_2. One now finds the next S on the basis of the new ∂_1 and ∂_2 and so on.

Remark 2.1. So far we assumed that the $\tau^{(1)}(x)$ and $\tau^{(2)}(x)$ are exponentially distributed. As we saw, the resulting basic model is a generalization of Richardson's growth model. It is well known that Richardson's growth model is a special case of first–passage percolation. We therefore describe here a corresponding generalization of first–passage percolation with different passage time distributions, depending on the number of occupied neighbors, but one in which the $\tau^{(i)}(x)$ do not need to be exponentially distributed. Instead we take $F^{(i)}$ as a general distribution on $[0, \infty)$ and take all $\tau^{(i)}(x)$ independent, now with $\tau^{(1)}(x)$ having distribution function $F^{(1)}(\varepsilon \cdot)$ and $\tau^{(2)}(x)$ having distribution function $F^{(2)}((1-\varepsilon)\cdot)$. The set of occupied sites at time $t, B(t)$, is defined in the same way as above. The basic model is obtained as a special case of this last model by taking $F^{(1)}$ and $F^{(2)}$ standard exponential distributions, while the classical first–passage percolation model corresponds to taking the limit $F^{(2)} \Rightarrow \delta_\infty$, that is $F^{(2)}(x) \to 0$ for each fixed x. We shall not prove anything about this more general model, but we believe that most of Theorems 1.1 and 1.2 will continue to hold provided the distributions $F^{(i)}$ have exponentially bounded tails, have a density $f^{(i)}$ on $[0, a]$ for some $a > 0$ and satisfy the following conditions:

$$0 < \liminf_{x \downarrow 0} f^{(i)}(x) \le \limsup_{x \downarrow 0} f^{(i)}(x) < \infty,$$

and for some $0 < C_1, C_2 < \infty$

$$\frac{F^{(i)}(x+y) - F^{(i)}(x-)}{1 - F^{(i)}(x-)} \le C_1 y, \tag{2.2}$$

$$\frac{1 - F^{(i)}(x+y-)}{1 - F^{(i)}(x-)} \le C_2[1 - F^{(i)}(y-)] \quad x, y \ge 0. \tag{2.3}$$

These properties come in when one must bound the rate at which a site becomes occupied, conditionally on the site not yet being occupied but having i neighbors occupied already for x time units. (2.2) ((2.3)) says that this conditioning does not drastically increase (repectively, decrease) the probability for a site to become occupied in the next y units of time (when compared to the unconditional probability). E.g. (2.2) prevents the process from exploding.

To carry out our proofs it is convenient to also define the analogous process $\eta^A_\cdot$ which has initial state

$$\eta^A_0 = 1\!\!1_A,$$

where A is any given finite subset of $\mathbb{Z}^d$. This process starts with all points in A occupied, rather than just the origin. It will be defined by means of the same $\tau^{(i)}(x)$ as the process $\eta_\cdot$. For this definition we identify our probability space with the space of all values of $\{\tau^{(i)}(x) : i = 1, 2, x \in \mathbb{Z}^d\}$. For a given sample point ω, *i.e.*, a specification of these variables, we can then define the processes $\eta^A_\cdot$ simultaneously for all finite $A \subset \mathbb{Z}^d$. η^A_t is defined in the same way as η_t above, except that we take $T(x) = 0$ for all $x \in A$ (rather than just for $x = \mathbf{0}$) and $B(0) = A$. We shall write η^x_t for $\eta^{\{x\}}_t$. The previous η_t is then the same as $\eta^{\mathbf{0}}_t$.

It is important for our arguments that this construction makes $\eta^A_\cdot$ into a strong Markov process with the jump rate from B to $B \cup y$ given by

$$\begin{aligned} &0 \text{ if } y \notin \partial B,\\ &\varepsilon \text{ if } y \in \partial_1 B,\\ &1 \text{ if } y \in \partial_2 B. \end{aligned}$$

Moreover, this process does not explode, as one can see for instance by comparison with the process in which each site in $\partial B(t)$ becomes occupied at rate 1. (Alternatively, one can use the fact that if t_k is the k–th jump time of the process, then the cardinality of $B(t_k)$ equals $|A| + k$ w.p. 1 (where $|A|$ stands for the cardinality of A), and given $B(t_k), t_{k+1} - t_k$ has an exponential distribution with mean at least

$$\frac{1}{|\partial B(t_k)|} \geq \frac{1}{2d(|A| + k)}.$$

Since $\sum(|A| + k)^{-1} = \infty$ our claims follows from Breiman (1968), Sects 15.5, 15.6.

A simple but very useful fact is that if C, D are finite subsets of $\mathbb{Z}^d$ for which $C \subset D$, then the simultaneously constructed processes $\eta^C_\cdot$ and $\eta^D_\cdot$ satisfy

$$\eta^C_t(x) \leq \eta^D_t(x) \text{ for all } t \geq 0, x \in \mathbb{Z}^d, \tag{2.4}$$

on each sample point for which

$$\begin{aligned} C(s) := \{x : \eta^C_s(x) = 1\} \text{ is finite for all } s \geq 0 \text{ and}&\\ \tau^{(i)}(x) < \infty \text{ for } i = 1, 2 \text{ and all } x \in \mathbb{Z}^d.& \end{aligned} \tag{2.5}$$

We shall not prove this, but the proof is very similar to the following lemma which holds even for the partially oriented model defined before Theorem 1.3. Note that the processes in both sides of (2.4) use the *same* $\tau^{(i)}(x)$ and that this is a deterministic lemma.

Lemma 2.1. *Let $C \subset \mathbb{Z}^d$ be finite and let the sample point be such that (2.5) holds. Then*

$$\eta^C_{t+s}(x) \geq \eta^{C(t)}_s(x), \quad s \geq 0, x \in \mathbb{Z}^d. \tag{2.6}$$

Proof. Fix $t \geq 0$ and assume that (2.6) fails. Then

$$S := \inf\{s : \eta^{C(t)}_s(x) = 1 \text{ and } \eta^C_{t+s}(x) = 0 \text{ for some } x\} < \infty.$$

Thus there exist $s_k \downarrow S, x_k \in \mathbb{Z}^d$ such that

$$\eta^{C(t)}_{s_k}(x_k) = 1, \eta^C_{t+s_k}(x_k) = 0.$$

It is not hard to see that if $\eta^{C(t)}_s(x) = 1$, then there exists a path $y_0, y_1, \dots, y_n$ on $\mathbb{Z}^d$ with $y_0 \in C(t), y_n = x$ and $\eta^{C(t)}_s(y_i) = 1$ for $0 \leq i \leq n$. We claim that this allows us to choose $d(x_k, C(t+S+1)) \leq 1$, where $d(x, C) := \inf\{||x - y||_\infty : y \in C\}$. To see this assume that $s_k < S + 1$ and $\eta^{C(t)}_{s_k}(x_k) = 1$ and $\eta^C_{t+s_k}(x_k) = 0$ but $d(x_k, C(t+S+1)) > 1$. Then choose a path $y_0 \in C(t), \dots, y_n = x$ from $C(t)$ to x for which $\eta^{C(t)}_{s_k}(y_i) = 1$ as above. If y_r is the first point on this path outside $C(t+s_k)$, and hence $\eta^C_{t+s_k}(y_r) = 0$, then we may replace x_k by y_r. Indeed, this y_r satisfies

$$d(y_r, C(t+S+1)) \leq d(y_r, C(t+s_k)) \leq 1,\ \eta^{C(t)}_{s_k}(y_r) = 1,\ \eta^C_{t+s_k}(y_r) = 0.$$

From now on we assume that $d(x_k, C(t+S+1)) \leq 1$. Since, by assumption, $C \subset C(t) \subset C(t+S+1)$ is a finite set, x_k takes on finitely many values only and there must exist an x such that $x_k = x$ for infinitely many k. Then the right continuity of $s \to \eta^{C(t)}_s(x)$ and of $s \to \eta^C_{t+s}(x)$ shows that

$$\eta^{C(t)}_S(x) = 1, \quad \eta^C_{t+S}(x) = 0. \tag{2.7}$$

On the other hand, the minimality of S implies that

$$\eta^{C(t)}_s(y) \leq \eta^C_{t+s}(y) \text{ for all } s < S \text{ and all } y \in \mathbb{Z}^d,$$

and in particular $\eta^{C(t)}_s(x) = 0$ for $s < S$, so that $T(x) \geq S$, where for the course of this proof we use $T(z)$ for the first time that z becomes occupied in the $\eta^{C(t)}_\cdot$ process. Similarly we shall use $V^{(i)}(z)$ for the first time i neighbors of z become occupied in the $\eta^{C(t)}_\cdot$ process. Thus

$$T(x) = S$$

and this means that

$$S = \min(V^{(1)}(x) + \tau^{(1)}(x), V^{(2)}(x) + \tau^{(2)}(x)).$$

For the sake of argument, let $S = V^{(2)}(x) + \tau^{(2)}(x) > V^{(2)}(x)$. But then $V^{(2)}(x) = T(z_2)$ for some neighbor z_2 of x with $T(z_2) < S$ and there exists another neighbor

z_1 of x with $T(z_1) \le T(z_2) < S$. Then $1 = \eta^{C(t)}_{T(z_j)}(z_j) \le \eta^C_{t+T(z_j)}(z_j)$, that is, z_j is already occupied at time $t + T(z_j)$ in the process $\eta^C_\cdot$. Finally then, at time $t + T(z_2) + \tau^{(2)}(x) = t + S, x$ must also be occupied in the $\eta^C_\cdot$ process (by (2.1)), or equivalently $\eta^C_{t+S}(x) = 1$, which contradicts (2.7). ■

We remark that the hypothesis (2.5) of Lemma 2.1 is fulfilled w.p. 1 whenever C is finite. Clearly $\tau^{(i)}(x) > 0$ w.p. 1, and the finiteness of $C(s)$ for $s \ge 0$ follows from the fact that $\eta^C_\cdot$ does w.p. 1 not explode.

One immediate consequence of (2.6) is the subadditivity relation

$$T(x+y) \le T(x) + T^x(x+y), \tag{2.8}$$

where $T(z)$ and $T^x(z)$ denote the first time that z becomes occupied in the process $\eta_\cdot = \eta^{\mathbf{0}}_\cdot$ and the process $\eta^x_\cdot$, respectively. This can be seen by taking $C = \{0\}, t = T(x), s = T^x(x+y)$ and by replacing x by $x+y$ in (2.6). Then $x \in C(T(x))$ and

$$\eta_{T(x)+s}(x+y) \ge \eta^{C(T(x))}_s(x+y) \ge \eta^x_s(x+y) = \eta^x_{T^x(x+y)}(x+y) = 1,$$

from which (2.8) follows. (2.8) of course implies (1.2) via Kingman's subadditive ergodic theorem, and the existence of the asymptotic shape $A(\varepsilon)$ as in (1.3) follows by Richardson's argument (cf. Durrett (1988), Ch. 1).

Repeated use will be made of the following simple lemma, which again holds for the partially oriented model.

Lemma 2.2. *Let $C \subset \mathbb{Z}^d$ be finite and $D \subset \mathbb{Z}^d$ arbitrary. Assume that $t_0 < \infty$ and $\delta \ge 0$ are such that*

$$\begin{aligned} &\mathbb{P}\{\eta^C_{t_0} \ge \mathbb{1}_D\} \\ &\quad = \mathbb{P}\{\textit{all of } D \textit{ is occupied at time } t_0 \textit{ by the process which starts with } C\} \ge \delta. \end{aligned}$$

Then

$$\mathbb{P}\{\eta^C_{kt_0} \ge \mathbb{1}_D\} \ge 1 - (1-\delta)^k, \quad k \ge 0.$$

Proof. We use induction on k. Let

$$\mathcal{F}_t = \sigma\{\eta^C_s : s \le t\}. \tag{2.9}$$

Then it suffices to show

$$\mathbb{P}\{\eta^C_{(k+1)t_0} \ge \mathbb{1}_D | \mathcal{F}_{kt_0}\} \ge \delta.$$

But this follows easily from the Markov property, since, with $C(s)$ as in (2.5),

$$\begin{aligned} \mathbb{P}\{\eta^C_{(k+1)t_0} \ge \mathbb{1} | \mathcal{F}_{kt_0}\} &= \mathbb{P}\{\eta^{C(kt_0)}_{t_0} \ge \mathbb{1}_D\} \\ &\ge \mathbb{P}\{\eta^C_{t_0} \ge \mathbb{1}_D\} \ge \delta. \end{aligned}$$

For the one but last inequality we used the fact that $C(kt_0) \supset C$ and (2.4). ■

Remark 2.2. The preceding lemma and Lemma 4.1 are the main steps which are nontrivial to generalize to the more general model in which the $\tau^{(i)}(x)$ are no longer exponentially distributed.

We now turn to the **proof of Theorem 1.1**. The lower bound on $v(\varepsilon, d)$ is the easier one to obtain. We first describe the argument in a somewhat informal way. The idea is that of considering a 'moving front' of the right size, embedded in the growth process. Consider the sets of sites

$$F_{n,m} := \{x = (x_1, x_2, \dots, x_n) \in \mathbb{Z}^d : x_1 = n \text{ and } 0 \le x_i \le m \text{ for } i = 2, \dots, d\}.$$

If at a certain time all the sites of $F_{n,m}$ are occupied, and all the sites of $F_{n+1,m}$ are vacant, then the rate at which a first site in $F_{n+1,m}$ becomes occupied is $\varepsilon(m+1)^{d-1}$. If we choose $m = \lfloor \varepsilon^{-\alpha} \rfloor$, then some site in $F_{n+1,m}$ will become occupied after a time of order $\varepsilon^{\alpha(d-1)-1} =: t_1$. On the other hand, once such a site is occupied, its neighbors in $F_{n+1,m}$ have at least two occupied neighbors, and, hence, are being occupied at rate 1. Once these sites are occupied, their neighbors in $F_{n+1,m}$ have at least two occupied neighbors, and, hence, are being occupied at rate 1, and so on. It is clear that $F_{n+1,m}$ should become completely occupied after an extra time of order $m = \varepsilon^{-\alpha} =: t_2$ beyond t_1. The order of the total time for $F_{n+1,m}$ to become completely occupied by the mechanism described above is minimized by making t_1 and t_2 be of the same order. This is accomplished by taking $\alpha = 1/d$. The total time for $F_{n+1,m}$ to become occupied then is of the order $\varepsilon^{-1/d}$. Repeating this reasoning for successive values of n, one sees why we used the term 'moving front' above. At each time there is a largest n for which $F_{n,m}$ is completely occupied, and the $F_{n,m}$ for this n is our moving front. This front moves with speed given by the inverse of the time needed for it to move one unit, *i.e.*, at a speed of order $\varepsilon^{1/d}$.

To transform the reasoning above into a rigorous argument, we will compare the growth model with a modified model in which each one of the sets $F_{n,m}$ can only start to become occupied after the set $F_{n-1,m}$ is fully occupied (except for the case $n = 0$, for which we use the original dynamics). For concreteness we suppose also that in this modified process the sites which do not belong to any of the sets $F_{n,m}$, $n = 0, 1, \dots$, stay vacant forever. We will call this modified process the 'front process'. It is easy to see how to couple the original growth model and the front process so that the former one has always a larger set of occupied sites than the latter. To obtain such a coupling we construct the front process using the same exponential clocks as used to construct the growth process. We will denote by $T^{\mathrm{f}}(x)$ the time when the site x is first occupied in the front process, and by $V^{\mathrm{f}(1)}(x)$ and $V^{\mathrm{f}(2)}(x)$ the times when the site x first has, respectively, 1 or 2 occupied neighbors in this process. The front process is constructed using the exponential clocks $\tau^{(1)}(x)$ and $\tau^{(2)}(x)$ by means of the following definitions. First, $T^{\mathrm{f}}(\mathbf{0}) = 0$ and if $x \notin \cup_{n \ge 0} F_{n,m}$ then $T^{\mathrm{f}}(x) = \infty$. For x in $F_{0,m}$ we define $T^{\mathrm{f}}(x)$ by using the original basic model restricted to $F_{0,m}$, that is, we use the growth process as if none of the vertices outside $F_{0,m}$ are present. Now, define inductively $\Theta_0 := 0$,

$$\Theta_n := \max\{T^{\mathrm{f}}(x) : x \in F_{n-1,m}\},$$

for $n > 0$, and for $x \in F_{n,m}$, $n \geq 0$,

$$T^{\mathrm{f}}(x) := \min\{\max\{V^{\mathrm{f}(1)}(x), \Theta_n\} + \tau^{(1)}(x), \max\{V^{\mathrm{f}(2)}(x), \Theta_n\} + \tau^{(2)}(x)\}.$$

We leave it to the reader to show (by arguments similar to those in Lemma 2.1) that indeed for all $x, T^{\mathrm{f}}(x) \geq T(x)$, as claimed.

Based on the informal computations above, we will choose $m = \lfloor \varepsilon^{-1/d} \rfloor$. We have to show that in the front process, the expectation of the random time $\Delta_n := \Theta_{n+1} - \Theta_n$, for the set $F_{n,m}$ to become fully occupied after the set $F_{n-1,m}$ became fully occupied, satisfies

$$\mathbb{E}(\Delta_n) < C_1 \varepsilon^{-1/d}, \tag{2.10}$$

for an appropriate constant C_1, for all $n > 0$ and all $\varepsilon > 0$. (This inequality is false for $n = 0$, but the obvious fact $\Delta_0 < \infty$ a.s. is enough for our purposes.) The lower bound in the theorem follows immediately from (2.10), the law of large numbers (*i.e.*, $\frac{1}{n}\Theta_n \to \mathbb{E}\Delta_1$) and the comparison between the original and the front processes (which yields $T(n\mathbf{e}_1) \leq \Theta_n$). To check (2.10) it is enough to show that for some C_2 and $\delta > 0$, independent of ε,

$$\mathbb{P}\{\Delta_n > C_2 \varepsilon^{-1/d}\} < 1 - \delta. \tag{2.11}$$

Indeed, (2.11) implies

$$\mathbb{P}\{\Delta_n > k C_2 \varepsilon^{-1/d}\} < (1 - \delta)^k, \quad k \geq 0, \tag{2.12}$$

by virtue of Lemma 2.2. To show (2.11), divide the random time interval Δ_n into the time $\Delta_n^{(1)}$ needed for a first site in $F_{n,m}$ to be occupied and the extra time $\Delta_n^{(2)}$ needed for that set to become fully occupied after one of its sites first became occupied. It is clear that $\Delta_n^{(1)}$ is an exponential random variable with mean $[(m+1)^{d-1}\varepsilon]^{-1} \leq \varepsilon^{-1/d}$ and that

$$\mathbb{P}\{\Delta_n^{(1)} > \varepsilon^{-1/d}\} \leq e^{-1}.$$

So all we need is to also show that for some constant C_3,

$$\mathbb{P}\{\Delta_n^{(2)} > C_3 \varepsilon^{-1/d}\} \leq 1/2. \tag{2.13}$$

In dimension $d = 2$, (2.13) follows immediately by comparison with a pair of rate 1 Poisson processes because of the argument given above. If E_z is the event that z is the first site in $F_{n,m}$ to become occupied, then on the event E_z with $z = (z_1, z_2)$, the difference between the times at which $(z_1, z_2 + k)$ and z become occupied is bounded by

$$\sum_{j=1}^{k} \tau^{(2)}(z_1, z_2 + j) ;$$

a similar estimate holds for the difference between the times at which $(z_1, z_2 - k)$ and z become occupied. In higher dimensions, the same sort of comparison with rate 1 Poisson processes can also be used to prove (2.13). Given $z \in F_{n,m}$, choose for each site $x \in F_{n,m}$ a sequence of sites $z = x(0), x(1), \dots, x(n) = x$ such that $||x(i) - x(i-1)||_1 = 1$ for $i = 1, \dots, n$ and $n = ||x - z||_1$. In particular $n \le dm$. In the front process, when the site $x(i)$ is occupied and $x(i+1)$ is not yet occupied, this latter site is being occupied at rate 1 (because it has the occupied neighbors x_i and $x_{i+1} - \mathbf{e}_1$. If we constructed the front process using the exponential clocks as indicated above we can specifically write

$$\mathbb{P}\{T^{\mathrm{f}}(x) > \Theta_n + \Delta_n^{(1)} + s | E_z\} \le \mathbb{P}\{\sum_{i=1}^{n} \tau^{(2)}(x(i)) > s\} \le \mathbb{P}\{Z < n\} \le \mathbb{P}\{Z < dm\},$$

where Z is a Poisson random variable with mean s. By taking $s = C_3\varepsilon^{-1/d}$, averaging over z, and using the standard large deviation estimate (A.1) (see the Appendix for formulae numbered (A.$*$)), one obtains

$$\begin{aligned}\mathbb{P}\{\Delta_n^{(2)} > C_3\varepsilon^{-1/d}\} &\le (m+1)^d \mathbb{P}\{Z < dm\} \\ &\le (\varepsilon^{-1/d}+1)^d C_0 (d\varepsilon^{-1/d}+1) e^{(d-C_3)\varepsilon^{-1/d}} \left(\frac{C_3}{d}\right)^{d\varepsilon^{-1/d}}.\end{aligned} \tag{2.14}$$

This implies (2.13) when C_3 is large enough, and finishes the proof of the lower bound in Theorem 1.1.

We turn now to the proof of the upper bound on $v(\varepsilon, d)$ in Theorem 1.1. First we define three concepts which are useful below: A site $x \in \mathbb{Z}^d$ is called a *slow site* if it becomes occupied at a moment in which it has exactly one occupied neighbor, *i.e.*, if $T(x) = V^{(1)}(x) + \tau^{(1)}(x)$. A *chain* in $\mathbb{Z}^d$ is a sequence of distinct sites $x(1), x(2), \dots, x(n)$ such that, for $i = 1, \dots, n-1$, $x(i)$ and $x(i+1)$ are neighbors. We say that $x(1)$ and $x(n)$ are the endpoints of the chain and that n is its length. A *chronological path* at time t is a chain of sites, for which $0 \le T(x(1)) < T(x(2)) < \cdots < T(x(n)) \le t$. We will say that this path goes from $x(1)$ to $x(n)$.

To prove the upper bound in Theorem 1.1, we will show that if a site is occupied at time t, then it has to be connected to the origin by a chronological path which contains (in a certain sense) many slow sites. Once this is realized, the proof is completed by estimating the probability of such an event. In order to be more precise about what we mean with 'many slow sites' we need the notation

$$H_j := \{x = (x_1, x_2, \dots, x_d) \in \mathbb{Z}^d : x_1 = j\}. \tag{2.15}$$

Lemma 2.3. *(Deterministic statement) For arbitrary $k > 0$, if the site $x \in H_k$ is occupied at time t, then there exists at time t a chronological path from the origin to x which contains at least one slow site in each one of the hyperplanes H_j, $j = 1, \dots, k$.*

Proof. We construct such a chronological path by moving backwards in time. Start from $x(n) = x$. We have to say how to find $x(i-1)$, once we have $x(j)$, $j = i, \dots, n$. It is clear that if a site which is not the origin is occupied at time t, then at least one of its neighbors was occupied before this site. To assure that in the path that we are constructing we actually obtain a slow site in each $H_j, j \leq k$, we use the following rule. Choose $x(i-1)$ from the neighbors of $x(i)$ which were occupied before $x(i)$, but if there is more then one such site never choose $x(i) - \mathbf{e}_1$ (otherwise the choice is arbitrary). In other words, only choose $x(i) - \mathbf{e}_1$ for $x(i-1)$ if forced to do so. This means that if $x(i-1) = x(i) - \mathbf{e}_1$, then $x(i)$ is a slow site. It is clear that, because only finitely many sites are occupied at any given time, and only the origin is occupied at time 0, the procedure just described produces a chronological path with the desired properties. (Note that n is determined only after the whole path has been constructed.) ∎

Suppose now that at time t there is an occupied site x in H_k with $k = \lfloor \lambda t \rfloor$ for some $\lambda > 0$. Then by Lemma 2.3 the event

$$\begin{aligned} F(t,\lambda) := \{&\text{at time } t \text{ there exists a chronological path from the origin} \\ &\text{to some point in } H_k \text{ with a slow site in each } H_j, 1 \leq j \leq k\} \end{aligned}$$

must occur. The proof of the upper bound in Theorem 1.1 will be completed once we show that there is a constant C_1 and an $\varepsilon_0 > 0$ such that if $\lambda = C_1 \varepsilon^{1/d}$ and $\varepsilon \leq \varepsilon_0$, then

$$\lim_{t \to \infty} \mathbb{P}\{F(t,\lambda)\} = 0. \tag{2.16}$$

In order to prove (2.16), we first observe that we can choose a constant C_2 so large that no chronological path present at time t is likely to be longer than $C_2 t$. More precisely, for suitable C_3, C_4

$$\begin{aligned} \mathbb{P}\{&\text{there exists a chronological path starting at the origin} \\ &\text{and of length } \geq C_2 t\} \leq C_3 e^{-C_4 t}. \end{aligned} \tag{2.17}$$

This is a well known fact, which follows from the observation that the number of chains of sites which have the origin as one endpoint and have length ℓ is bounded above by $2d(2d-1)^{\ell-1}$, while the probability that any given one of these chains is a chronological path present at time t is bounded above by $\mathbb{P}\{Z \geq \ell - 1\}$, where Z is a Poisson random variable with mean t. The standard large deviations estimate (A.2) for Poisson random variables implies then that the probability of having any chronological path with length at least $C_2 t$ present at time t is bounded above by

$$\sum_{\ell \geq C_2 t} 2d(2d-1)^{\ell-1} \mathbb{P}\{Z \geq \ell - 1\} \leq 2d \sum_{\ell \geq C_2 t - 1} (2d-1)^{\ell} e^{-\ell(\log(C_2)-1)}.$$

By taking C_2 large, this upper bound tends to 0 exponentially fast as $t \to \infty$.

The problem of proving (2.16) is now reduced to showing that

$$\lim_{t \to \infty} \mathbb{P}\{G(t,\lambda,C_2)\} = 0, \tag{2.18}$$

where $G(t,\lambda,C_2)$ is the event that at time t there is a chronological path from the origin to some point in H_k with length not larger than $C_2 t$ and which contains at least one slow site in each one of the hyperplanes H_i, $i = 1, \ldots, k$ (recall that $k = \lfloor \lambda t \rfloor$).

In order to estimate the probability that $G(t, \lambda, C_2)$ occurs, we basically need an upper bound for the number of choices for the set of slow sites which appears in the definition of this event. For a sample point at which $G(t, \lambda, C_2)$ occurs, let $S = \{y(1), \dots, y(k)\}$ be a set of sites on a chronological path from $\mathbf{0}$ to a point in H_k at time t, such that $y(j) \in H_j$, and such that $y(j)$ is a slow site. There may be several possible choices for S, but this will not influence our argument. One may make the choice of S unique by ordering the collection of possible sets S lexicographically and by picking for S the first possible one in this ordering. Define for each j, $z(j) = z(j; S) = y(j) - y(j-1)$, where $y(0) = \mathbf{0}$. The components of these vectors will be identified by the usual notation $z(j) = (z_1(j), \dots, z_d(j))$. The set S is completely specified by the sequence $(z(j) : j = 1, \dots, k)$, or, equivalently, by the sequences $(z_i(j) : j = 1, \dots, k)$, $i = 2, \dots, d$ (while $z_1(j) = 1$ for each j).

Since $G(t, \lambda, C_2)$ cannot occur if $\lambda > C_2$ we suppose $\lambda \le C_2$. If we define

$$K_i := \sum_{j=1}^{k} |z_i(j)|, \tag{2.19}$$

then for any chronological path of length $\le C_2 t$ and corresponding set S, we must have

$$K_i \le C_2 t. \tag{2.20}$$

Let $S = \{y(1), \dots, y(k)\}$ be a set of sites such that $y(j) \in H_j$, $j = 1, \dots, k$ and such that (2.20) is satisfied. The number of ways in which the sequence $(z_i(j) : j = 1, \dots, k)$ can be chosen (for fixed i) is bounded above by

$$\sum_{K_i=1}^{\lfloor C_2 t \rfloor} 2^k \binom{K_i + k - 1}{k-1} \le C_2 t 2^k \binom{\lfloor C_2 t \rfloor + k}{k} \le C_3 t 2^k \left(\frac{2eC_2 t}{k} \right)^k,$$

for some constant C_3. The factor 2^k comes from the choices of the signs of the $z_i(j)$, $j = 1, \dots, k$, while the binomial coefficient comes from the number of ways in which the $|z_i(j)|$, $j = 1, \dots, k$, can be chosen, given the constraint (2.20) (see Feller (1968), Sect. II.5) . The final inequality above is a simple consequence of Stirling's formula and is derived in the Appendix, as (A.3); here we are also using the fact that $k \le \lambda t \le C_2 t$.

For $M(t, \lambda, C_2) :=$ the number of sets S which can arise, we now have the bound

$$M(t, \lambda, C_2) \le \left(C_3 t 2^k \left(\frac{2eC_2 t}{k} \right)^k \right)^{d-1} \le C_4 t^{d-1} \left(\frac{C_5}{\lambda} \right)^{k(d-1)}, \tag{2.21}$$

for some constants C_4 and C_5.

From the construction of the growth model using exponential clocks, one can see that the following holds. Given a set of sites S, of cardinality k, the probability that there is at time t a chronological path which goes through all these sites and that all sites in S are slow, is bounded above by

$$\mathbb{P}\left\{ \sum_{y \in S} \tau^{(1)}(y) \le t \right\} = \mathbb{P}\{Z' \ge k\},$$

where Z' is a Poisson random variable with mean $t\varepsilon$. Using again the standard large deviation estimate (A.2), we obtain, for $\lambda > \varepsilon$,

$$\mathbb{P}\{Z' \geq k\} \leq \left(\frac{t\varepsilon}{k}\right)^k e^k. \tag{2.22}$$

Combining this inequality with (2.21) and using $k = \lfloor \lambda t \rfloor$, we obtain for large t

$$\mathbb{P}\{G(t,\lambda,C_2)\} \leq M(t,\lambda,C_2)\mathbb{P}\{Z' \geq k\} \leq C_6 t^{d-1}\left(\frac{C_7\varepsilon}{\lambda^d}\right)^{\lambda t-1}, \tag{2.23}$$

for some constants C_6 and C_7. This inequality implies (2.18) when $\lambda = C_1\varepsilon^{1/d}$, and C_1 is chosen large enough. This completes the proof of the upper bound in Theorem 1.1.

3. Proof of Theorem 1.2.

First we introduce some terminology. In arbitrary dimension we shall use the term *rectangle* for a product of intervals, thus extending the usual meaning of a rectangle in two dimensions. When considering the lattice $\mathbb{Z}^d$, a rectangle will mean a product of lattice intervals. When considering the asymptotic shape $A(\varepsilon)$, we will be dealing with compact convex subsets of $\mathbb{R}^d$ which have many symmetries. It is natural to use the name *radius* of such a set for the smallest value of ℓ for which the set is contained in $Q(\ell)$, defined by (1.4). Equivalently, the radius of a set A equals $\max\{||x||_\infty : x \in A\}$. Below, we shall refer informally to the *two–neighbors mechanism.* By this we mean the phenomenon of vacant sites which have at least 2 occupied neighbors becoming occupied at rate 1.

Having tried to convince the reader, in Remark 1.2 in the introduction, that Theorem 1.2 is less obvious than it might seem at first sight, we will argue next that it should nevertheless be true. The rigorous proof will be a direct adaptation of a reasoning that we first present informally. We know that for fixed ε there exists an asymptotic shape, in the sense that (1.3) holds. So after a long time t, the fattened occupied region, $\overline{\eta}_t$, should be close to the set $tA(\varepsilon)$. Because of Theorem 1.1, the linear size of this set in the coordinate directions should be of order $r_t := C_3 t\varepsilon^{1/d}$ (here ε is fixed and t tends to ∞). In other words, at a large time t the coordinate axes should be occupied up to a distance from the origin of order r_t, but there should be no occupied vertices on the coordinate axes much further out than r_t. It is easy to see that $A(\varepsilon)$ is convex, compact and invariant under permutations and signchanges of the coordinates (compare Kesten (1986), pp. 158-160). As in Cox and Durrett (1981), proof of Theorem 5, it follows from this that $A(\varepsilon)$ is contained in the cube $Q(\ell_\varepsilon)$, where $\ell_\varepsilon = \max\{x : x\mathbf{e}_1 \in A(\varepsilon)\}$. That is, the furthest occupied points (in the $||\ ||_\infty$ sense) are roughly along the coordinate axes, and for any $\delta > 0, \overline{\eta}_t$ will eventually be contained in $Q((1+\delta)r_t)$. Wait now an extra time $C_4 t\varepsilon^{1/d}$, were C_4 is large. If ε is small, this extra time is short compared to t. Therefore, at time $t' = t(1 + C_4\varepsilon^{1/d})$ the occupied region should not be much bigger than the dilation by a factor close to 1 of the occupied region at time t. The fractional increment in the radius should actually be of order $C_4 t\varepsilon^{1/d} C_3\varepsilon^{1/d}/C_3 t\varepsilon^{1/d} = C_4\varepsilon^{1/d}$. The crucial step in the argument now is that during this extra time interval, the two–neighbors mechanism is likely to have filled up the cube of radius r_t (provided C_4 is large enough; this step has to be proven uniformly in ε). Before we argue for this last statement we observe that it leads to the desired conclusion, since then we have at time t' a shape which is larger than the cube of radius r_t, but smaller than a cube of radius $(1+\delta)r_{t'}$. Because t' can be made arbitrarily large by taking t large, what we see at time t' must be close to the asymptotic shape. The asymptotic shape $A(\varepsilon)$ must therefore be close to a cube.

The occupation of the cube of radius $r_t = C_3 t\varepsilon^{1/d}$ by the two-neighbors mechanism in a time $t' - t = C_4 t\varepsilon^{1/d}$, should occur as claimed above, if C_4 is large, because we are starting from a configuration in which each coordinate axis is occupied essentially up to a distance r_t of the origin. If time were discrete and the two-neighbor mechanism completely deterministic (at each unit of time, sites which have two occupied neighbors would become occupied), then this claim would be trivially true: indeed, if at time t each of the coordinate axes is occupied till distance r, then at time $t+s$ all sites in $\{x : ||x||_1 \leq s\}$ would be occupied, for $s \leq r$. In the continuous time setting that we are considering the proof of the same statement is not so trivial, and if one tries to directly adapt the argument indicated for the discrete time case, it will not work (for the time needed to fill up the cube one would obtain a useless upper bound of the order $r_t \log(r_t)$). We will state and prove next a precise version of what we need. After this is done, we will complete the proof of Theorem 1.2 along the lines sketched in the previous paragraph.

At no extra cost, we will state and prove the next lemma in a form which is slightly stronger than the one we use, but which clarifies its content better. First we introduce a new comparison process, which will be called *continuous time oriented bootstrap percolation model* or just CTOBPM (see Aizenman and Lebowitz (1988) and Schonmann (1992) to understand the motivation for the name).

Define the oriented neighborhood of a site x by

$$\mathcal{N}_{\mathrm{or}}(x) = \{-\mathbf{e}_i : 1 \leq i \leq d\} + x.$$

In the CTOBPM, a site which is vacant becomes occupied at rate 1 if and only if at least two sites in its oriented neighborhood are occupied. Otherwise the rate of becoming occupied is 0. As usual in growth models, occupied sites never become vacant. We will be concerned with this process started from an occupied set S at time 0, and will denote by ζ_t^S or $\zeta_t(S)$ the random set of occupied sites at time t. Note that in the CTOBPM sites become occupied slower than in the completely oriented model of Theorem 1.3 with $\nu = d$, because in the CTOBPM no site can become occupied before at least *two* of its neighbors are occupied. In particular ζ_t^S will be contained in the occupied set at time t for the basic process starting with S, *i.e.* $\zeta_t^S \subset \{x : \eta_t^S(x) = 1\}$. We will consider initial sets of the form

$$S(n_1, \ldots, n_d) := \cup_{i=1}^d \{k\mathbf{e}_i : 0 \leq k \leq n_i\}.$$

To state the lemma we need also the following notation for rectangles

$$R(n_1, \ldots, n_d) := \prod_{i=1}^d [0, n_i].$$

Lemma 3.1. *For the CTOBPM, there exist positive finite constants C_1, C_2 and C_3 such that if $n = n_1 + \cdots + n_d$, then*

$$\mathbb{P}\{R(n_1, \dots, n_d) \not\subset \zeta_{C_1 n}\big(S(n_1, \dots, n_d)\big)\} \leq C_2 \exp(-C_3 n).$$

Proof. We construct the CTOBPM using the mean 1 exponential clocks $\tau^{(3)}(x) := \min\{\tau^{(1)}(x), \tau^{(2)}(x)\}$. We will use the notation $T^{\mathrm{BP}}(x)$ for the time when x is first occupied and $V^{\mathrm{BP}}(x)$ for the first time when in $\mathcal{N}_{\mathrm{or}}(x)$ there are at least two occupied sites. The process $(\zeta_t^S)_{t \geq 0}$ is constructed by setting $T^{\mathrm{BP}}(x) = 0$ if $x \in S$, and for other values of x, inductively

$$T^{\mathrm{BP}}(x) = \min\{V^{\mathrm{BP}}(x) + \tau^{(3)}(x)\}.$$

For $x \in R(n_1, \dots, n_d)$ define $\mathcal{P}_x$ as the set of oriented paths from the origin to x, that is, the set of sequences of sites $0 = x(0), x(1), \dots, x(k) = x$ such that $x(i-1) \in \mathcal{N}_{\mathrm{or}}(x(i))$ for $i = 1, \dots, k$. We further define

$$\mathcal{P}_{(n_1, \dots, n_d)} = \cup_{x \in R(n_1, \dots, n_d)} \mathcal{P}_x.$$

The following inequality is the crucial step in the proof. Suppose that initially the occupied set is $S(n_1, \dots, n_d)$. Then for every $x \in R(n_1, \dots, n_d)$,

$$T^{\mathrm{BP}}(x) \leq \max_{\pi \in \mathcal{P}_x} \sum_{z \in \pi} \tau^{(3)}(z) \tag{3.1}$$

This inequality will be proven by induction. First we introduce some terminology. A site in $R(n_1, \dots, n_d)$ will be said to be in the k–th class if exactly k of its coordinates are different from 0. Clearly (3.1) is true if x is in the 0–th or 1–st class for then $x \in S$. Suppose now that $k \geq 2$ and that (3.1) is true for sites in the 0–th to $(k-1)$–th classes and also for sites $x = (x_1, \dots, x_d)$ in the k–th class which have $||x||_1 = x_1 + \cdots + x_d < j$. We want to show that then it will also hold for sites y in the k–th class which have $||y||_1 = j$. To see this, observe that for such a y at least 2 coordinates are different from 0, and hence at least 2 sites in $\mathcal{N}_{\mathrm{or}}(y)$ are in the set of sites for which we supposed that (3.1) holds. Call these sites y' and y'' (the choice is arbitrary if there are more than 2 sites). From the construction of the model with exponential clocks we have

$$T^{\mathrm{BP}}(y) \leq \max\{T^{\mathrm{BP}}(y'), T^{\mathrm{BP}}(y'')\} + \tau^{(3)}(y).$$

This inequality and the induction hypothesis imply that (3.1) also holds for y (even if we restrict the maximum there to paths which pass through either y' or y''). This completes the proof of (3.1).

Using (3.1) we have

$$\begin{aligned}
&\mathbb{P}\{R(n_1, \dots, n_d) \not\subset \zeta_{C_1 n}\big(S(n_1, \dots, n_d)\big)\} \\
&= \mathbb{P}\{T^{\mathrm{BP}}(x) > C_1 n \text{ for some } x \in R(n_1, \dots, n_d)\} \\
&\leq \mathbb{P}\{\max_{\pi \in \mathcal{P}_{(n_1, \dots, n_d)}} \sum_{z \in \pi} \tau^{(3)}(z) > C_1 n\}.
\end{aligned}$$

We can now use again a well known type of estimate, which was also used in the last section. First one observes that the cardinality of $\mathcal{P}_{(n_1,\dots,n_d)}$ is bounded above by d^n, where, as defined before, $n = n_1 + \cdots + n_d$. On the other hand, for each $\pi \in \mathcal{P}_{(n_1,\dots,n_d)}$, $\mathbb{P}\{\sum_{z\in\pi} \tau^{(3)}(z) > C_1 n\} = \mathbb{P}\{Z < \text{ length of } \pi\} \le \mathbb{P}\{Z < n\}$, where Z is a Poisson random variable with mean $C_1 n$. Using (A.1), we obtain now

$$\mathbb{P}\{R(n_1,\dots,n_d) \not\subset \zeta_{C_1 n}\big(S(n_1,\dots,n_d)\big)\} \le d^n C_0 (n+1) e^{n - C_1 n} (C_1)^n,$$

which decays exponentially with n if C_1 is large. ∎

Recall now the definition (1.4) and define $\ell(\varepsilon)$ as the radius of $\varepsilon^{-1/d} A(\varepsilon)$. Tautologically,

$$\varepsilon^{-1/d} A(\varepsilon) \ \subset \ Q(\ell(\varepsilon)).$$

From Theorem 1.1 we know that for some $0 < C_1 \le C_2 < \infty$,

$$C_1 \le \ell(\varepsilon) \le C_2.$$

The proof of Theorem 1.2 is reduced now to showing that for arbitrary $\delta > 0$ there exists ε_0, such that for all $\varepsilon < \varepsilon_0$,

$$(1-\delta) Q(\ell(\varepsilon)) \ \subset \ \varepsilon^{-1/d} A(\varepsilon). \tag{3.2}$$

For each fixed ε, (1.3) holds and hence, given $\phi > 0$, we have for large t

$$\mathbb{P}\{(1-\phi) s A(\varepsilon) \subset \overline{\eta}_s \subset (1+\phi) s A(\varepsilon) \ \text{ for all } s \ge t\} \ge 1/2.$$

Set $t' = t'(t, \varepsilon, C_4) := t + C_4 \varepsilon^{1/d} t$. Then, from the above, we obtain

$$\mathbb{P}\{(1-\phi) t A(\varepsilon) \subset \overline{\eta}_t \text{ and } \overline{\eta}_{t'} \subset (1+\phi) t' A(\varepsilon)\} \ge 1/2. \tag{3.3}$$

Because $\ell(\varepsilon)$ is the radius of $\varepsilon^{-1/d} A(\varepsilon)$, comparison between the growth model and the CTOBPM and the symmetries of the growth model yield

$$\begin{aligned}
&\mathbb{P}\{(1-\phi) t A(\varepsilon) \subset \overline{\eta}_t \text{ and } (1-\phi) t \varepsilon^{1/d} Q(\ell(\varepsilon)) \not\subset \overline{\eta}_{t'}\} \\
&\le 2^d \mathbb{P}\{R\big((1-\phi) t \varepsilon^{1/d} l(\varepsilon), \dots, (1-\phi) t \varepsilon^{1/d} l(\varepsilon)\big) \\
&\qquad\qquad \not\subset \zeta_{C_4 \varepsilon^{1/d} t}\big(S((1-\phi) t \varepsilon^{1/d} l(\varepsilon), \dots, (1-\phi) t \varepsilon^{1/d} l(\varepsilon)\big)\}.
\end{aligned}$$

By virtue of Lemma 3.1 and the fact that $\ell(\varepsilon) \le C_2 < \infty$ this inequality implies that for large C_4 (independent of ε)

$$\mathbb{P}\{(1-\phi) t A(\varepsilon) \subset \overline{\eta}_t \text{ and } (1-\phi) t \varepsilon^{1/d} Q(\ell(\varepsilon)) \not\subset \overline{\eta}_{t'}\} \to 0, \tag{3.4}$$

as $t \to \infty$, provided $C_4 \ge C_1 C_2$, where C_1 is as in Lemma 3.1. Combining (3.3) and (3.4) we conclude that there are points in the sample space for which we can find t so that

$$(1-\phi) t \varepsilon^{1/d} Q(\ell(\varepsilon)) \ \subset \ \overline{\eta}_{t'} \ \subset \ (1+\phi) t' A(\varepsilon). \tag{3.5}$$

Therefore for each $\varepsilon > 0$ and $\phi > 0$,

$$Q(\ell(\varepsilon)) \subset \varepsilon^{-1/d} A(\varepsilon) \frac{t'}{t} \frac{1+\phi}{1-\phi} = \varepsilon^{-1/d} A(\varepsilon) (1 + C_4 \varepsilon^{1/d}) \frac{1+\phi}{1-\phi}.$$

Because C_4 does not depend on ϕ or ε, and ϕ is arbitrary, this leads to the desired conclusion (3.2), and completes the proof of Theorem 1.2. Actually it leads to the semi–explicit expression

$$\frac{1}{1 + C_4 \varepsilon^{1/d}} Q(\ell(\varepsilon)) \subset \varepsilon^{-1/d} A(\varepsilon).$$

The constant C_4 taken as $C_2 C_1$ (see (3.4)).

4. Proof of Theorem 1.3.

First we shall prove the estimates for the point–to–point speeds v and w, which are easy. Fix $0 \le \nu \le d$ and consider the model with the last ν coordinates oriented. In this model, if x becomes occupied at the time $T(x)$, then at least one of its neighbors $y \in x + \mathcal{N}_\nu$ must have been occupied. Accordingly, we change the definition of chronological path somewhat from the definition used for Lemma 2.3. Now a chain $x(1), \dots, x(n)$ will be called a chronological path at time t if $0 \le T(x(1)) < \cdots < T(x(n)) \le t$ and $x(i) \in x(i+1) + \mathcal{N}_\nu,\ 1 \le i \le n$. With this definition the statement and proof of Lemma 2.3 carry over unchanged to the model with the last ν coordinates oriented. In particular, a vertex x is occupied at time t if and only if there exists a chronological path at t from the origin to x.

Now observe that if $\mathbf{0} = x(1), \dots, x(\ell) = n\mathbf{e}_j$ is a chronological path from the origin to the point $n\mathbf{e}_j$ on the j–th coordinate axis, then $x(i+1) - x(i) = \pm \mathbf{e}_{r(i)}$ for some $r(i)$ (because $x(i+1)$ and $x(i)$ are adjacent on $\mathbb{Z}^d$). If $d - \nu + 1 \le r(i) \le d$, then we must even have $x(i+1) - x(i) = \mathbf{e}_{r(i)}$, because the last ν coordinates can only increase along a chronological path. This implies that for any $s \in [d - \nu + 1, d]$ with $s \neq j$, no steps at all can be taken in the s–th direction (otherwise we would have $x_s(\ell) = s$–th coordinate of $n\mathbf{e}_j > s$–th coordinate of $\mathbf{0} = 0$). Therefore the presence of the coordinates numbered $d - \nu + 1$ to d, other than j, has no influence on the rate at which $n\mathbf{e}_j$ becomes occupied. When we remove the coordinates numbered $d - \nu + 1$ to d, other than the j–th coordinate, we obtain (1.5) and (1.7).

(1.6) is now immediate from (1.5) and Theorem 1.1. Also the upper bound in (1.8) follows from (1.5) and Theorem 1.1, because any speed in a partially oriented model cannot be greater than the same speed in the totally unoriented model of the same dimension. As for the lower bound in (1.8), this still follows as in Theorem 1.1 from the moving front construction, because in this construction occupation was allowed to move in the positive first coordinate direction only (from $F_{n-1,m}$ to $F_{n,m}$). Thus, apart from an interchange between the first and last coordinate, the proof of Theorem 1.1 still applies to $w(\varepsilon, d - \nu + 1, 1)$.

Next we turn to proving the right hand inequalities in (1.9) and (1.10). These proofs are still essentially the same as for the upper bound in Theorem 1.1. We begin with (1.10). Consider the events

$$\begin{aligned} L(t,\lambda) = \{&\text{at time } t \text{ there exists an occupied vertex} \\ &x = (x_1, \dots, x_d) \text{ with } x_{d-\nu+1} = \lfloor \lambda t \rfloor, x_j \le \lfloor \lambda t \rfloor, d - \nu + 2 \le j \le d\}. \end{aligned}$$

and

$$\begin{aligned} \widetilde{G}(t, \lambda, C_2) = \{&\text{at time } t \text{ there exists a chronological path of length} \\ &\le C_2 t \text{ from } \mathbf{0} \text{ to some point } x = (x_1, \dots, x_d) \text{ with} \\ &x_{d-\nu+1} = \lfloor \lambda t \rfloor, x_j \le \lfloor \lambda t \rfloor, d - \nu + 2 \le j \le d, \text{and at} \\ &\text{least one slow point in each } \widetilde{H}_i, i = 1, \dots, \lfloor \lambda t \rfloor\}, \end{aligned}$$

where

$$\widetilde{H}_i = \{x = (x_1, \dots, x_d) : x_{d-\nu+1} = i\}.$$

Then as in the proof of Theorem 1.1, with C_2 as in (2.17),

$$\mathbb{P}\{L(t,\lambda)\} \le C_3 e^{-C_4 t} + \mathbb{P}\{\widetilde{G}(t,\lambda,C_2)\}.$$

If $\widetilde{G}(t,\lambda,C_2)$ occurs, let $x(1),\dots,x(n)$ be a chronological path at time t from $\mathbf{0}$ to a point in $\widetilde{H}_{\lfloor\lambda t\rfloor}$ with the properties stated in the definition of $\widetilde{G}(t,\lambda,C_2)$. As in the proof of Theorem 1.1, let $S=\{y(1),\dots,y(\lfloor\lambda t\rfloor)\}$ be a set of slow sites on this path, one in each $\widetilde{H}_i, 1\le i\le\lfloor\lambda t\rfloor$. With $z(j)=y(j)-y(j-1)$, as before, (2.20) must still hold for all $1\le i\le d$. However, for $d-\nu+1\le i\le d$ we even have the stronger requirement

$$z_i(j)\ge 0 \text{ and } K_i=\sum_{j=1}^{\lfloor\lambda t\rfloor} z_i(j)\le x_i(n)\le\lfloor\lambda t\rfloor. \tag{4.1}$$

Therefore, for $d-\nu+2\le i\le d$, the number of ways in which the sequence $\{z_i(j), 1\le j\le\lfloor\lambda t\rfloor\}$ can be chosen is bounded above by

$$\sum_{K_i=1}^{\lfloor\lambda t\rfloor}\binom{K_i+\lfloor\lambda t\rfloor-1}{\lfloor\lambda t\rfloor-1}\le\lambda t\binom{2\lfloor\lambda t\rfloor}{\lfloor\lambda t\rfloor}\le C_3 t C_4^{\lfloor\lambda t\rfloor}$$

(by (A.3)). Consequently, if we again denote the number of ways in which S can be chosen by $M(t,\lambda,C_2)$, then (2.21) can be replaced by

$$\begin{aligned} M(t,\lambda,C_2) &\le \left(C_3 t 2^{\lfloor\lambda t\rfloor}\left(\frac{2eC_2t}{\lfloor\lambda t\rfloor}\right)^{\lfloor\lambda t\rfloor}\right)^{d-\nu}\left(C_3 t C_4^{\lfloor\lambda t\rfloor}\right)^{\nu-1}\\ &\le C_5 t^{d-1}\Big(\frac{C_6}{\lambda^{d-\nu}}\Big)^{\lambda t}. \end{aligned}$$

As in (2.23) this yields, again with Z' a Poisson variable with mean $t\varepsilon$,

$$\mathbb{P}\{\widetilde{G}(t,\lambda,C_2)\}\le M(t,\lambda,C_2)\mathbb{P}\{Z'\ge\lfloor\lambda t\rfloor\}\le C_5t^{d-1}\Big(\frac{C_7\varepsilon}{\lambda^{d-\nu+1}}\Big)^{\lambda t}.$$

It follows that for

$$\lambda > C_8\varepsilon^{1/(d-\nu+1)},$$

with C_8 sufficiently large,

$$\mathbb{P}\{L(t,\lambda)\}\le C_9e^{-C_{10}t}. \tag{4.2}$$

By symmetry in the last ν coordinates we even have

$$\mathbb{P}\{\text{for arbitrarily large } t \text{ there exists at time } t \text{ an occupied vertex } x \\ \text{with } \max_{d-\nu+1\le j\le d} x_j\ge C_8\varepsilon^{1/(d-\nu+1)}t\}=0.$$

The right hand inequality of (1.10) follows.

Next we prove the right hand inequality in (1.9). This time we define

$$\widehat{L}(t,\lambda) = \{\text{at time } t \text{ there exists an occupied vertex } x = (x_1,\dots,x_d) \\ \text{with } x_1 = \lfloor \lambda t \rfloor \text{ and } x_i \le C_8 \varepsilon^{1/(d-\nu+1)} t \text{ for } d-\nu+1 \le i \le d\}$$

and (with H_j as in (2.15))

$$\widehat{G}(t,\lambda,C_2) = \{\text{at time } t \text{ there exists a chronological path of length } \le C_2 t \text{ from } \mathbf{0} \\ \text{to some point } x \text{ in } H_{\lfloor \lambda t \rfloor}, \text{ with at least one slow site in each } H_j, \\ 1 \le j \le \lfloor \lambda t \rfloor, \text{ and with } 0 \le x_j \le C_8 \varepsilon^{1/(d-\nu+1)} t, d-\nu+1 \le j \le d\}.$$

Again we take C_2 so large that (2.17) holds. On the event $\widehat{L}(t,\lambda)$ one can construct, as in Lemma 2.3, a chronological path at time t from $\mathbf{0}$ to x which has a slow point in each $H_j, 1 \le j \le \lfloor \lambda t \rfloor$. Moreover, by (4.2) we know that

$$\mathbb{P}\{\text{at time } t \text{ there exists an occupied vertex } x \text{ with} \\ x_i > C_8 \varepsilon^{1/(d-\nu+1)} t \text{ for some } d-\nu+1 \le i \le d\} \le d C_9 e^{-C_{10} t}.$$

From these relations it follows as in the proof of Theorem 1.1 that

$$\begin{aligned} &\mathbb{P}\{\text{at time } t \text{ there exists an occupied point } x \text{ with } x_1 \ge \lfloor \lambda t \rfloor\} \\ &\le \mathbb{P}\{\widehat{L}(t,\lambda)\} + d C_9 e^{-C_{10} t} \\ &\le \mathbb{P}\{\widehat{G}(t,\lambda,C_2)\} + C_3 e^{-C_4 t} + d C_9 e^{-C_{10} t}. \end{aligned}$$

It therefore suffices to prove that

$$\mathbb{P}\{\widehat{G}(t,\lambda,C_2)\} \to 0 \text{ exponentially fast} \tag{4.3}$$

for any $\lambda \ge C_{11} \varepsilon^{\frac{d+1}{d(d-\nu+1)}}$ for C_{11} sufficiently large. It even suffices to prove this for $\lambda = \left\lfloor C_{11} \varepsilon^{\frac{d+1}{d(d-\nu+1)}} \right\rfloor$.

(4.3) can now be proven by imitating the end of the proof of Theorem 1.1. If $\widehat{G}(t,\lambda,C_2)$ occurs, and $x(1),\dots,x(n)$ is a chronological path with the properties listed in the definition of $\widehat{G}(t,\lambda,C_2)$, then let $S = \{y(1),\dots,y(\lfloor \lambda t \rfloor)\}$ be a set of slow points on this path, one from each $H_j, 1 \le j \le \lfloor \lambda t \rfloor$. With $z(j) = y(j) - y(j-1)$ as before, (2.20) must again hold for all $1 \le i \le d-\nu$. For $d-\nu+1 \le i \le d$ the relation

$$z_i(j) \ge 0 \text{ and } K_i = \sum_{j=1}^{\lfloor \lambda t \rfloor} z_i(j) \le C_8 \varepsilon^{1/(d-\nu+1)} t$$

must hold (compare (4.1)). It follows that this time

$$M(t,\lambda,C_2) \le \left(C_3 t 2^{\lfloor \lambda t \rfloor} \left(\frac{2e C_2 t}{\lfloor \lambda t \rfloor} \right)^{\lfloor \lambda t \rfloor} \right)^{d-\nu-1} \left(\binom{C_8 \varepsilon^{1/(d-\nu+1)} t + \lfloor \lambda t \rfloor}{\lfloor \lambda t \rfloor} \right)^{\nu}.$$

We take $\lambda = \left\lfloor C_{11}\varepsilon^{\frac{d+1}{d(d-\nu+1)}} \right\rfloor$ and for brevity write $\gamma = C_8\varepsilon^{1/(d-\nu+1)}$. Without loss of generality we take ε so small that $\lambda \le \gamma$. By (A.3) we then have

$$M(t,\lambda,C_2) \le C_{12}t^{d-\nu-1}\left(\frac{C_{13}}{\lambda}\right)^{(d-\nu-1)\lambda t}\left(\frac{2e\gamma}{\lambda}\right)^{\nu\lambda t}.$$

As in (2.22) and (2.23) we finally obtain

$$\mathbb{P}\{\widehat{G}(t,\lambda,C_2)\} \le M(t,\lambda,C_2)\left(\frac{te\varepsilon}{\lfloor\lambda t\rfloor}\right)^{\lfloor\lambda t\rfloor} \le C_{14}t^{d-\nu-1}\left(\frac{C_{15}\gamma^\nu\varepsilon}{\lambda^d}\right)^{\lambda t}.$$

Substituting our values for λ and γ, we see that for sufficiently large C_{11} (independent of ε) this tends to 0 exponentially fast in t (for fixed ε), so that (4.3) follows.

As for the left hand inequality in (1.10) this is immediate from the left hand inequality in (1.8) and the fact that the point–to–hyperplane speed in any direction is at least as large as the point–to–point speed in the same direction.

Finally, for the lower bound in (1.9) we shall use the following lemma.

Lemma 4.1. *Let $C \subset \mathbb{Z}^d$ be a finite set which contains $\mathbf{0}$ and define $\mathcal{F}_t$ as in (2.9). Fix $1 \le r \le d$. Let Θ be an $\mathcal{F}_t$–stopping time and assume that w.p. 1*

$$\eta_\Theta^C \ge 1\!\!1_{C+a}$$

for some (random) vector $\overline{a} = (a_1,\dots,a_d)$ with

$$a_r = 1 \quad \text{and} \quad a_j \ge 0, \quad d-\nu+1 \le j \le d. \tag{4.4}$$

Then

$$\liminf_{n\to\infty}\frac{n}{\sigma(n,r,\varepsilon,\nu)} \ge \frac{1}{\mathbb{E}\Theta}. \tag{4.5}$$

Proof. Of course (4.5) is equivalent to

$$\limsup\frac{\sigma(n,r,\varepsilon,\nu)}{n} \le \mathbb{E}\Theta \text{ w.p. 1}. \tag{4.6}$$

Note that the r–th component of $\overline{a}$ equals 1, so that at time Θ, the occupied set contains a copy of C, shifted one unit in the r–th direction (and shifted a_j units in the j–th direction for $j \ne r$). One can then expect that the occupied set will reach the hyperplane $\{x_r = n\}$ after n time intervals which are stochastically 'similar' to Θ (actually they will be stochastically smaller than Θ). This is the resaon for (4.6).

Now to the details. Note that C will be completely occupied after some finite random time κ. Define

$\sigma^C(n,r,\varepsilon,\nu) =$ first time a point in the hyperplane
$\{x : x_r = n\}$ is occupied in the $\eta_\cdot^C$–process .

Then it follows from Lemma 2.1 that

$$\sigma(n,r,\varepsilon,\nu) \le \kappa + \sigma^C(n,r,\varepsilon,\nu).$$

It therefore suffices to prove (4.6) with σ replaced by σ^C. Define

$$\begin{aligned}
&\Theta_0 = 0,\\
&\Theta_1 = \Theta, \quad a(1) = \overline{a},\\
&\Theta_{\ell+1} = \inf\{t > \Theta_\ell : \eta_t^C \ge 1\!\!1_{C+\overline{a}(\ell+1)} \text{ for some}\\
&\qquad \text{vector } \overline{a}(\ell+1) \text{ with } a_r(\ell+1) = \ell+1, a_j(\ell+1) \ge a_j(\ell), d-\nu+1 \le j \le d\}.
\end{aligned}$$

Note that $\overline{a}(1)$, and at the later stages $\overline{a}(\ell)$, may not be uniquely determined. Any $\mathcal{F}_{\Theta_\ell}$–measurable choice for $\overline{a}(\ell)$ will do for our purposes, though. Since $\overline{a}(n) \in C + \overline{a}(n)$ has r–th coordinate equal to n, we have

$$\sigma^C(n,r,\varepsilon,\nu) \le \Theta_n = \sum_{\ell=0}^{n-1} (\Theta_{\ell+1} - \Theta_\ell). \tag{4.7}$$

Furthermore by the strong Markov property applied to Θ_ℓ

$$\begin{aligned}
&\mathbb{P}\{\Theta_{\ell+1} - \Theta_\ell > s | \mathcal{F}_{\Theta_\ell}\}\\
&= \mathbb{P}\{\eta^C_{\Theta_\ell + s} \text{ does not contain any set of the form } C + \overline{a}(\ell+1) | \mathcal{F}_{\Theta_\ell}\}\\
&= \mathbb{P}\{\eta_s^{C(\Theta_\ell)} \text{ does not contain any set of the form } C + \overline{a}(\ell+1)\},
\end{aligned}$$

where $C(\Theta_\ell)$ is defined by (2.5) with $s = \Theta_\ell$ and $\overline{a}(\ell+1)$ has to have the properties listed in the definition of $\Theta_{\ell+1}$. But $C(\Theta_\ell) \supset C + \overline{a}(\ell)$, so that by (2.4) the last probability is at most

$$\begin{aligned}
&\mathbb{P}\{\eta_s^{C+\overline{a}(\ell)} \text{ does not contain any set of the form } C + \overline{a}(\ell+1)\}\\
&= \mathbb{P}\{\eta_s^C \text{ does not contain any set of the form } C + \overline{a}(\ell+1) - \overline{a}(\ell)\}\\
&\text{(by translation invariance)} = \mathbb{P}\{\Theta > s\}.
\end{aligned}$$

In summary, we have

$$\mathbb{P}\{\Theta_{\ell+1} - \Theta_\ell > s | \mathcal{F}_{\Theta_\ell}\} \le \mathbb{P}\{\Theta > s\}.$$

We can therefore couple the $\Theta_{\ell+1} - \Theta_\ell, \ell \ge 0$, with a sequence $\rho_\ell, \ell \ge 0$, of i.i.d. copies of Θ so that

$$\Theta_n \le \sum_{\ell=0}^{n-1} \rho_\ell, \quad n \ge 1.$$

This coupling is fairly standard; for a more general coupling result see Kamae, Krengel and O'Brien (1977). Together with (4.7) and the strong law of large numbers this proves

$$\limsup_{n\to\infty} \frac{\sigma(n,r,\varepsilon,\nu)}{n} \le \limsup_{n\to\infty} \frac{1}{n} \sum_{\ell=0}^{n-1} \rho_\ell = \mathbb{E}\Theta. \qquad \blacksquare$$

To prove the lower bound in (1.9) we apply Lemma 4.1 with

$$C = [0, L]^{d-\nu} \times \{0\}^{\nu},$$

where

$$M = \left\lfloor \varepsilon^{-\frac{1}{d(d-\nu+1)}} \right\rfloor \text{ and } L = M \left\lfloor \varepsilon^{-1/(d-\nu+1)} \right\rfloor \sim \varepsilon^{-\frac{d+1}{d(d-\nu+1)}}.$$

We define

$$\Theta = \inf\{t \geq 0 : \eta_t^C \geq 1\!\!1_{C+\overline{a}} \text{ for some vector } \overline{a} \text{ satisfying (4.4)}\}.$$

$\mathbb{E}\Theta$ is estimated in two stages. First we estimate $\mathbb{E}\Theta'$, where

$$\Theta' = \inf\{t \geq 0 : \eta_t^C \geq 1\!\!1_{[0,L]^{d-\nu} \times [0,M]^{\nu}}\}.$$

For any vector $\overline{j} = (j_1, \dots, j_{d-\nu}) \in [0, M-1]^{d-\nu}$ define

$$D = D(\overline{j}) = \prod_{i=1}^{d-\nu} [j_i \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor, j_{i+1} \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor] \times [0, M]^{\nu}.$$

For the time being we shall suppress the $\overline{j}$ in the notation.

To estimate $\mathbb{E}\Theta'$ we use the moving front method of the beginning of the proof of Theorem 1.1. Define for $\overline{b} = \{b_{d-\nu+1}, \dots, b_{\ell}\}$

$$E(\overline{b}) = E(\overline{b}, \overline{j}) = \prod_{i=1}^{d-\nu} [j_i \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor, j_{i+1} \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor] \times \{b_{d-\nu+1}, \dots, b_{\ell}\}$$

and fix $d - \nu + 1 \leq r \leq d$. Then there exists a $C_1 < \infty$ and $\delta > 0$ (independent of $\overline{b}, \overline{j}$ and r) for which

$$\begin{aligned}
&\mathbb{P}\{\eta^{E(\overline{b})}_{C_1 \varepsilon^{-1/(d-\nu+1)}} \geq 1\!\!1_{E(\overline{b}+\mathbf{e}_r)}\} \\
&= \mathbb{P}\{\text{at time } C_1 \varepsilon^{-1/(d-\nu+1)} \text{ the set} \\
&\quad \prod_{i=1}^{d-\nu} [j_i \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor, j_{i+1} \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor] \times \{b_{d-\nu+1}, \dots, b_{r-1}, b_r + 1, b_{r+1}, \dots, b_{\ell}\} \\
&\qquad \text{is fully occupied by the } \eta^{E(\overline{b})}_{\cdot}\text{–process}\} \\
&\geq \delta.
\end{aligned}$$

This is simply estimate (2.11) in the dimensions $1, \dots, d-\nu, r$; the dimensions numbered $d - \nu + 1, \dots, d$ other than r play no role in this estimate. By Lemma 2.2 we then also have

$$\begin{aligned}
&\mathbb{P}\{\prod_{i=1}^{d-\nu} [j_i \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor, j_{i+1} \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor] \times \{b_{d-\nu+1}, \dots, b_{r-1}, b_r + 1, b_{r+1}, \dots, b_{\ell}\} \\
&\text{ is not yet fully occupied by the } \eta^{E(\overline{b})}_{\cdot}\text{–process at time } kC_1 \varepsilon^{-1/(d-\nu+1)}\} \\
&\leq (1-\delta)^k. \qquad (4.8)
\end{aligned}$$

Now we essentially repeat the argument leading to (2.14). Let $\overline{c} = (c_{d-\nu+1}, \dots, c_d) \in [0, M]^\nu$ and consider the oriented path on $\mathbb{Z}^\nu$ from $\mathbf{0}$ to $\overline{c}$ which first goes form $(0, \dots, 0)$ to $(c_{d-\nu+1}, 0, \dots, 0)$ by $c_{d-\nu+1}$ steps in the positive $(d-\nu+1)$–th coordinate direction, next goes from $(c_{d-\nu+1}, 0, \dots, 0)$ to $(c_{d-\nu+1}, c_{d-\nu+2}, 0, \dots, 0)$ in $c_{d-\nu+2}$ steps in the $(d-\nu+2)$–th coordinate direction $\dots$ and finally from $(c_{d-\nu+1}, \dots, c_{d-1}, 0)$ to $\overline{c}$ in c_d steps in the d–th coordinate direction. The total number of steps in this path is $n := \sum_{d-\nu+1}^{d} c_i \le \nu M$. We write this path as $\overline{b}(0) = \mathbf{0}, \overline{b}(1), \dots, \overline{b}(n) = \overline{c}$ and define

$$\begin{aligned}
\kappa_0 &= \kappa_0(\overline{c}) = 0\\
\kappa_{i+1} &= \kappa_{i+1}(\overline{c}) = \inf\{t \ge \kappa_i : E(\overline{b}(i+1))\text{is occupied by } \eta_t^C\}\\
&= \inf\{t \ge \kappa_i : \eta_t^C \ge 1\!\!1_{E(\overline{b}(i+1))}\}.
\end{aligned}$$

Note that

$$\eta_0^C = 1\!\!1_C \ge 1\!\!1_{E(\overline{b}(0))},$$

because for $\overline{j} \in [0, M-1]^{d-\nu}$,

$$E(b(0)) = \prod_{i=1}^{d-\nu} [j_i \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor, j_{i+1} \lfloor \varepsilon^{-1/(d-\nu+1)} \rfloor] \times \{0\}^\nu \subset C.$$

Also, by definition, for $i \ge 1, \eta_{\kappa_i}^C \ge 1\!\!1_{E(\overline{b}(i))}$ or $C(\kappa_i) \supset E(\overline{b}(i))$, with $C(s)$ as in (2.5). Again by the strong Markov property,

$$\begin{aligned}
&\mathbb{P}\{\kappa_{i+1} - \kappa_i > k\varepsilon^{-1/(d-\nu+1)} | \mathcal{F}_{\kappa_i}\}\\
&= \mathbb{P}\{\eta^C_{\kappa_i + k\varepsilon^{-1/(d-\nu+1)}} \text{ does not occupy } E(\overline{b}(i+1)) | \mathcal{F}_{\kappa_i}\}\\
&= \mathbb{P}\{\eta^{C(\kappa_i)}_{k\varepsilon^{-1/(d-\nu+1)}} \ge 1\!\!1_{E(\overline{b}(i+1))} \text{ fails}\}\\
&\le \mathbb{P}\{\eta^{E(\overline{b}(i))}_{k\varepsilon^{-1/(d-\nu+1)}} \ge 1\!\!1_{E(\overline{b}(i+1))} \text{ fails}\}\\
&\le (1-\delta)^k \quad \text{(by (4.8))}.
\end{aligned}$$

Moreover, at time κ_n, the whole set $E(b(n)) = E(\overline{c})$ is occupied. Therefore

$$\begin{aligned}
&\mathbb{P}\{E(\overline{c}) \text{ is not yet occupied at time } \ell\varepsilon^{-1/(d-\nu+1)} \text{ by the } \eta_\cdot^C\text{–process}\\
&\le \mathbb{P}\{\kappa_n > \ell\varepsilon^{-1/(d-\nu+1)}\}\\
&\le \mathbb{P}\{\sum_0^{n-1} \lambda_i > \ell\},
\end{aligned}$$

where $\lambda_0, \lambda_1, \dots$ are i.i.d. with

$$\mathbb{P}\{\lambda_i > k\} = (1-\delta)^k, \quad k = 0, 1, \dots.$$

A standard large deviation estimate shows that for large C_2, some constant C_3, and all $\ell \geq C_2 n$,

$$\begin{aligned}&\mathbb{P}\{\sum_0^{n-1} \lambda_i > \ell\}\\ &\leq \inf_{s\geq 0} e^{-\ell s}\left(\mathbb{E}e^{s\lambda_0}\right)^n\\ &\leq e^{-C_3\ell}.\end{aligned}$$

Therefore, for each $\overline{c} \in [0,M]^\nu$ and $\ell \geq C_2\nu M$,

$$\begin{aligned}&\mathbb{P}\{E(\overline{c}) \text{ is not yet occupied at time } \ell\varepsilon^{-1/(d-\nu+1)}\}\\ &\leq \mathbb{P}\{\sum_0^{n-1}\lambda_i > \ell\} \leq e^{-C_3\ell}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{P}\{D(\overline{j}) \text{ is not yet occupied at time } \ell\varepsilon^{-1/(d-\nu+1)} \text{ by the } \eta^C_\cdot\text{–process}\}\\ &= \mathbb{P}\{\prod_{i=1}^{d-\nu}[j_i\lfloor\varepsilon^{-1/(d-\nu+1)}\rfloor, j_{i+1}\lfloor\varepsilon^{-1/(d-\nu+1)}\rfloor] \times [0,M]^\nu \text{ is not yet}\\ &\text{occupied at time } \ell\varepsilon^{-1/(d-\nu+1)}\text{by the } \eta^C_\cdot\text{–process}\}\\ &\leq (M+1)^\nu e^{-C_3\ell}.\end{aligned}$$

(The factor $(M+1)^\nu$ represents the number of $\overline{c}$ in $[0,M]^\nu$.) Finally, if $D(\overline{j})$ is occupied for all $\overline{j} \in [0, M-1]^{d-\nu}$, then also $C \times [0,M]^\nu$ is occupied. Therefore for $\ell \geq C_2\nu M$ and ε so small that $(M+1)^d \leq e^{C_2C_3\nu M/2}$ we have

$$\begin{aligned}&\mathbb{P}\{\Theta' > \ell\varepsilon^{-1/(d-\nu+1)}\}\\ &= \mathbb{P}\{D(\overline{j}) \text{ is not yet occupied at time } \ell\varepsilon^{-1/(d-\nu+1)} \text{ for some } \overline{j} \in [0,M-1]^{d-\nu}\}\\ &\leq M^{d-\nu}(M+1)^\nu e^{-C_3\ell} \leq e^{C_2C_3\nu M/2}e^{-C_3\ell} \leq e^{-C_3\ell/2}.\end{aligned}$$

It folows that

$$\mathbb{E}\Theta' \leq C_4 M\varepsilon^{-1/(d-\nu+1)} \leq C_4\varepsilon^{-\frac{d+1}{d(d-\nu+1)}} \tag{4.9}$$

for some constant C_4. This completes the first stage.

For the second stage we set

$$F = [0,L]^{d-\nu} \times [0,M]^\nu$$

and consider the process $\eta^F_\cdot$. We define

$$\Theta'' = \inf\{t \geq 0 : \eta^F_t \geq \mathbb{1}_{C+\overline{a}} \text{ for some vector } a \text{ satisfying (4.4)}\}.$$

We then have

$$\begin{aligned}
&\mathbb{P}\{\Theta - \Theta' > s | \mathcal{F}_{\Theta'}\} \\
&= \mathbb{P}\{\eta^C_{\Theta'+s} \text{ does not fully occupy any } C + \overline{a} \text{ for an } \overline{a} \text{ satisfying (4.4)} | \mathcal{F}_{\Theta'}\} \\
&= \mathbb{P}\{\eta^{C(\Theta')}_s \text{ does not fully occupy any } C + \overline{a} \text{ for an } \overline{a} \text{ satisfying (4.4)}\} \\
&\le \mathbb{P}\{\eta^F_s \text{ does not fully occupy any } C + \overline{a} \text{ for an } \overline{a} \text{ satisfying (4.4)}\} \\
&\text{(because } F \subset C(\Theta')) = \mathbb{P}\{\Theta'' > s\}.
\end{aligned} \tag{4.10}$$

Consequently

$$\mathbb{E}(\Theta - \Theta') \le \mathbb{E}\Theta'',$$

Assume now that we also prove for some constant C_5 that

$$\mathbb{E}\Theta'' \le C_5 \varepsilon^{-\frac{d+1}{d(d-\nu+1)}}. \tag{4.11}$$

Then we obtain from (4.9) that

$$\mathbb{E}\Theta \le (C_4 + C_5)\varepsilon^{-\frac{d+1}{d(d-\nu+1)}},$$

so that by Lemma 4.1

$$\liminf_{n\to\infty} \frac{n}{\sigma(n,r,\varepsilon,\nu)} \ge (C_4 + C_5)^{-1} \varepsilon^{\frac{d+1}{d(d-\nu+1)}}.$$

Thus we only have to prove (4.11) to complete the proof of (1.9).

To prove (4.11) we follow the argument for (2.14). We shall be dealing with vectors $\overline{a} = (a_1, \dots, a_d)$ which satisfy

$$\begin{aligned}
a_1 = L+1,\ 0 \le a_i \le L \text{ for } 2 \le i \le d-\nu,& \\
\text{and } 0 \le a_i \le M \text{ for } d-\nu+1 \le i \le d.&
\end{aligned} \tag{4.12}$$

We define

$$\begin{aligned}
\Delta^{(1)} = \inf\{t \ge 0 : \overline{a} = (L+1, a_2, \dots, a_d) \text{ is occupied by } \eta^F_t& \\
\text{for some } \overline{a} \text{ satisfying (4.12)}\},&
\end{aligned}$$

and take for a^* some $\mathcal{F}_{\Delta^{(1)}}$–measurable choice of an $\overline{a}$ which satisfies (4.12) and which is occupied at time $\Delta^{(1)}$. Thus $\eta^F_{\Delta^{(1)}}(a^*) = 1$. In addition we define

$$\begin{aligned}
\Delta^{(2)} &= \inf\{t \ge 0 : [1, L+1] \times [0,L]^{d-\nu-1} \times \{a^*_{d-\nu+1}, \dots, a^*_d\} \\
&\qquad \text{is occupied by } \eta^F_{\Delta^{(1)}+t}\} \\
&= \inf\{t \ge 0 : C + \{1, 0, \dots, 0, a^*_{d-\nu+1}, \dots, a^*_d\} \text{ is occupied by } \eta^F_{\Delta^{(1)}+t}\}.
\end{aligned}$$

Then $\Delta \le \Delta^{(1)} + \Delta^{(2)}$ and hence

$$\mathbb{E}\Delta \le \mathbb{E}\Delta^{(1)} + \mathbb{E}\Delta^{(2)}.$$

Now $\Delta^{(1)}$ is bounded by the minimum of $(L{+}1)^{d-\nu-1}(M{+}1)^{\nu}$ exponential variables of mean $1/\varepsilon$, so that

$$\mathbb{E}\Delta^{(1)} \le (L+1)^{-(d-\nu-1)}(M+1)^{-\nu}\frac{1}{\varepsilon} \le C_6\varepsilon^{-\frac{d+1}{d(d-\nu+1)}}. \tag{4.13}$$

Moreover, as in (4.10),

$$\begin{aligned}&\mathbb{P}\{\Delta^{(2)} \le s|\mathcal{F}_{\Delta^{(1)}}\}\\ &\qquad\qquad \ge \mathbb{P}\{C + \{1,0,\dots,0,a^*_{d-\nu+1},\dots,a^*_d\} \text{ is occupied by } \eta^{F\cup a^*}_{\Delta^{(1)}+s}\}.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}\Delta^{(2)} \le \max_{\overline{a}} \mathbb{E}\{&\text{time for } C + (1,0,\dots,0,a^*_{d-\nu+1},\dots,a^*_d)\\ &\text{to become fully occupied by } \eta^{F\cup\{a^*\}}\}.\end{aligned} \tag{4.14}$$

Now the argument which gave (2.14) shows that uniformly in a^* whose coordinates satisfy (4.12),

$$\begin{aligned}&\mathbb{P}\{C + (1,0,\dots,0,a^*_{d-\nu+1},\dots,a^*_d) \text{ is not fully occupied by } \eta^{F\cup\{a^*\}}_s\}\\ &\qquad\qquad \le (L+1)^{d-\nu-1}\mathbb{P}\{Z(s) < (d-\nu-1)L\},\end{aligned} \tag{4.15}$$

where $Z(s)$ is a Poisson variable with mean s. Indeed, $\eta^{F\cup\{a^*\}}$ starts with

$$(C + (0,\dots,0,a^*_{d-\nu+1},\dots,a^*_d) \cup \{a^*\} \subset F \cup \{a^*\}$$

already occupied. To occupy a point $\overline{c} = (L+1, c_2, \dots, c_{d-\nu}, a^*_{d-\nu+1},\dots,a^*_d)$ with $0 \le c_i \le L$ at time s, it suffices to occupy a path from $a^* = (L+1, a^*_2,\dots,a^*_d)$ to $\overline{c}$ in $\{L+1\} \times [0,L]^{d-\nu-1} \times \{a^*_{d-\nu+1},\dots,a^*_d\}$ by time s. But the points on any such path are successively occupied at rate 1 by the two–neighbors mechanism. Finally, (4.14) and (4.15) give for suitable C_7 – C_9

$$\begin{aligned}\mathbb{E}\Delta^{(2)} &= \int_0^\infty \mathbb{P}\{\Delta^{(2)} > s\}\\ &\le C_7L + (L+1)^{d-\nu-1}\int_{C_7L}^\infty \mathbb{P}\{Z(s) < (d-\nu-1)L\}\\ &\le C_7L + (L+1)^{d-\nu-1}\int_{C_7L}^\infty e^{(d-\nu-1)L}\mathbb{E}e^{-Z(s)}\,ds\\ &\le C_8L \le C_9\varepsilon^{-\frac{d+1}{d(d-\nu+1)}}.\end{aligned}$$

Together with (4.13) this proves (4.11) and completes the proof of the left hand inequality in (1.9).

Appendix.

In this appendix we prove a few elementary inequalities used in the paper. One of the ingredients in the derivations will be Stirling's formula, $n! \sim (2\pi n)^{(1/2)} n^n e^{-n}$ as $n \to \infty$.

First we derive two upper bounds for the large deviations of a Poisson random variable, Z, with mean μ. If $\ell < \mu$, then $e^{-\mu}\mu^k/k!$ is an increasing function of $k \in \{0, 1, \ldots, \ell\}$. Hence

$$\mathbb{P}\{Z \le \ell\} = \sum_{k=0}^{\ell} e^{-\mu}\frac{\mu^k}{k!} \le (\ell+1)e^{-\mu}\frac{\mu^\ell}{\ell!} \le C_0(\ell+1)e^{\ell-\mu}\left(\frac{\mu}{\ell}\right)^\ell, \tag{A.1}$$

where in the last step we used Stirling's formula, and C_0 is an appropriate constant.

If Z is still as above, but now $\ell \ge \mu$, then

$$\begin{aligned}\mathbb{P}\{Z \ge \ell\} = \sum_{i\ge\ell} e^{-\mu}\frac{\mu^i}{i!} \le \mu^\ell \sum_{i\ge\ell}\frac{\mu^{i-\ell}}{i!} \le \mu^\ell \sum_{i\ge\ell}\frac{\ell^{i-\ell}}{i!} \\ \le \left(\frac{\mu}{\ell}\right)^\ell \sum_{i\ge\ell}\frac{\ell^i}{i!} \le \left(\frac{\mu}{\ell}\right)^\ell e^\ell = e^{-\left(\log\left(\frac{\ell}{\mu}\right)-1\right)\ell}.\end{aligned} \tag{A.2}$$

Next we derive an inequality involving the binomial coefficients. First observe that for arbitrary strictly positive integers A and B,

$$\binom{A}{B} \le \frac{A^B}{B!} \le C_0\frac{A^B}{B^{1/2}B^Be^{-B}} \le C_0\left(\frac{eA}{B}\right)^B,$$

where C_0 is an appropriate constant. Therefore, if $0 < B \le A$, then

$$\binom{A+B}{B} \le C_0\left(\frac{e(A+B)}{B}\right)^B \le C_0\left(\frac{2eA}{B}\right)^B. \tag{A.3}$$

References

M. Aizenman and J. L. Lebowitz (1988), *Metastability effects in bootstrap percolation*, J. Phys. A **21**, 3801–3813.

L. Breiman (1968), *Probability*, Addison–Wesley Publ. Co..

J. T. Cox and R. Durrett (1981), *Some limit theorems for percolation processes with necessary and sufficient conditions*, Ann. Probab. **9**, 583–603.

R. Durrett (1988), *Lecture notes on interacting particle systems and percolation*, Wadsworth & Brooks/Cole Publ. Co..

R. Durrett and T. M. Liggett (1981), *The shape of the limit set in Richardson's growth model*, Ann. Probab. **9**, 186–193.

M. Eden (1961), *A two dimensional growth process*, Proc. 4th Berkeley Symposium Math. Statist. Probab., Vol IV, Univ. of California Press, pp. 223-239.

W. Feller (1968), *An Introduction to Probability Theory and Its Applications*, 3–rd ed., John Wiley & Sons.

T. Kamae, U. Krengel and G. L. O'Brien (1977), *Stochastic inequalities on partially ordered spaces*, Ann. Probab. **5**, 899-912.

H. Kesten (1986), *Aspects of first passage percolation*, Lecture Notes in Mathematics, vol. 1180 (P.L. Hennequin, ed.), Springer–Verlag, pp. 125-264.

D. Richardson (1973), *Random growth in a tesselation*, Proc. Camb. Phil. Soc. **74**, 515–528.

R. H. Schonmann (1992), *On the behavior of some cellular automata related to bootstrap percolation*, Ann. Probab. **20**, 174–193.

R. H. Schonmann (1994), *Theorems and conjectures on the droplet driven relaxation of stochastic Ising models*, in *Probability theory of spatial disorder and phase transition*, G. Grimmett, ed., Kluwer Publ. Co.

K. Schürger (1979), *On the asymptotic geometrical behaviour of a class of contact interaction processes with a monotone infection rate*, Z. Wahrsch. verw. Geb. **48**, 35-48.

H. Kesten
Department of Mathematics
Cornell University
Ithaca, NY 14853

R.H.Schonmann
Mathematics Department
University of California at Los Angeles
Los Angeles, CA 90024