The Annals of Probability
2012, Vol. 40, No. 2, 743–764
DOI: 10.1214/10-AOP631

SUBLINEARITY OF THE TRAVEL-TIME VARIANCE FOR DEPENDENT FIRST-PASSAGE PERCOLATION

BY JACOB VAN DEN BERG AND DEMETER KISS[1]

CWI and VU University Amsterdam, and CWI

Let E be the set of edges of the d-dimensional cubic lattice $\mathbb{Z}^d$, with $d \geq 2$, and let $t(e), e \in E$, be nonnegative values. The passage time from a vertex v to a vertex w is defined as $\inf_{\pi : v \to w} \sum_{e \in \pi} t(e)$, where the infimum is over all paths π from v to w, and the sum is over all edges e of π.

Benjamini, Kalai and Schramm [2] proved that if the $t(e)$'s are i.i.d. two-valued positive random variables, the variance of the passage time from the vertex 0 to a vertex v is sublinear in the distance from 0 to v. This result was extended to a large class of independent, continuously distributed t-variables by Benaïm and Rossignol [1].

We extend the result by Benjamini, Kalai and Schramm in a very different direction, namely to a large class of models where the $t(e)$'s are dependent. This class includes, among other interesting cases, a model studied by Higuchi and Zhang [9], where the passage time corresponds with the minimal number of sign changes in a subcritical "Ising landscape."

1. Introduction and statement of results. Consider, for $d \geq 2$, the d-dimensional lattice $\mathbb{Z}^d$. Let $\mathbb{E}$ denote the set of edges of the lattice, and let $t(e), e \in \mathbb{E}$, be nonnegative real values. A path from a vertex v to a vertex w is an alternating sequence of vertices and edges

$$v_0 = v, e_1, v_1, e_2, \ldots, v_{n-1}, e_n, v_n = w,$$

where each e_i is an edge between the vertices v_{i-1} and v_i, $1 \leq i \leq n$. To indicate that e is an edge of a path π, we often write, with some abuse of notation, $e \in \pi$.

If $v = (v_1, \ldots, v_d)$ is vertex, we use the notation $|v|$ for $\sum_{i=1}^{d} |v_i|$. The (graph) distance $d(v, w)$ between vertices v and w is defined as $|v - w|$. The vertex $(0, \ldots, 0)$ will be denoted by 0.

The passage time of a path π is defined as

$$T(\pi) = \sum_{e \in \pi} t(e). \tag{1}$$

Received July 2010; revised October 2010.

[1]Supported by NWO.

MSC2010 subject classifications. Primary 60K35; secondary 82B43.

Key words and phrases. First-passage percolation, influence results, greedy lattice animals, Ising model.

The passage time (or travel time) $T(v, w)$ from a vertex v to a vertex w is defined as

$$T(v, w) = \inf_{\pi : v \to w} T(\pi),$$

where the infimum is over all paths π from v to w.

Analogous to the above described bond version, there is a natural site version of these notions: in the site version, the t-variables are assigned to the vertices instead of the edges. In the definition of $T(\pi)$, the right-hand side in (1) is then replaced by its analog where the sum is over all vertices of π. There seems to be no "fundamental" difference between the bond and the site version.

An important subject of study in first-passage percolation is the asymptotic behavior of $T(0, v)$ and it fluctuations, when $|v|$ is large and the $t(e)$'s are random variables. It is believed that, for a large class of distributions of the $t(e)$'s, the variance of $T(0, v)$ is of order $|v|^{2/3}$. However, this has only been proved for a special case in a modified (oriented) version of the model [11]. Apart from this, the best upper bounds obtained for the variance before 2003 were linear in $|v|$ [13]. See Section 1 of [2] for more background and references.

Benjamini, Kalai and Schramm [2] showed that if the $t(e)$'s are i.i.d. random variables taking values a and b, $b \geq a > 0$, then the variance of $T(0, v)$ is sublinear in the distance from 0 to v. More precisely, they showed the following theorem.

THEOREM 1.1 (Benjamini, Kalai and Schramm [2]). *Let $b \geq a > 0$. If the $(t(e), e \in \mathbb{E})$ are i.i.d. random variables taking values in $\{a, b\}$, then there is a constant $C > 0$ such that, for all v with $|v| \geq 2$,*

$$\operatorname{Var}(T(0, v)) \leq C \frac{|v|}{\log|v|}. \tag{2}$$

Benaïm and Rossignol [1] extended this result to a large class of i.i.d. t-variables with a continuous distribution, and also proved concentration results. See also [5].

We give a generalization of Theorem 1.1 in a very different direction, namely to a large class of *dependent* t-variables. The description of this class, and the statement of our general results are given in Section 1.4.

Using our general results, we show in particular that (2) holds for the $\{a, b\}$-valued Ising model with $0 < a < b$ and inverse temperature $\beta < \beta_c$. By $\{a, b\}$-valued Ising model, we mean the model that is simply obtained from the ordinary, $\{-1, +1\}$-valued, Ising model by replacing -1 by a and $+1$ by b. The precise definition of the Ising model and the statement of this result is given in Section 1.1.

We also study, as a particular case of our general results, a different Ising-like first-passage percolation model: consider an "ordinary" Ising model (with signs -1 and $+1$), with parameters $\beta < \beta_c$ and with external field h satisfying certain conditions. Now define the passage time $T(v, w)$ between two vertices v and w as the minimum number of sign changes needed to travel from v to w. Higuchi and

Zhang [9] proved, for $d = 2$, a concentration result for this model. This concentration result implies an upper bound for the variance that is (a "logarithmic-like" factor) *larger* than linear. We show from our general framework that the sublinear bound (2) holds (see Theorem 1.5).

The last special case we mention explicitly is that where the collection of t-variables is a finite-valued Markov random field which satisfies a high-noise condition studied by Häggström and Steif (see [7]). Again it follows from our general results that the sublinear bound (2) holds (see Theorem 1.4).

The general organization of the paper is as follows: in the next three subsections, we give precise definitions and statements concerning the special results mentioned above. Then, in Section 1.4, we state our main, more general results, Theorems 1.6 and 1.7.

In Section 2, we prove the special cases (Theorems 1.2, 1.4 and 1.5) from Theorems 1.6 and 1.7.

In Section 3, we present the main ingredients for the proofs of our general results: an inequality by Talagrand (and its extension to multiple-valued random variables), a very general "randomization tool" of Benjamini, Kalai and Schramm, and a result on greedy lattice animals by Martin [15].

In Section 4, we first give a very brief informal sketch of the proof of Theorem 1.6 (pointing out the extra problems that arise, compared with the i.i.d. case in [2]), followed by a formal, detailed proof.

The proof of Theorem 1.7 is very similar to that of Theorem 1.6. This is explained in Section 5.

1.1. *The case where the t-variables have an $\{a, b\}$-valued Ising distribution.* Recall that the Ising model (with inverse temperature β and external field h) on a countably infinite, locally finite graph G is defined as follows. First some notation: we write $v \sim w$ to indicate that two vertices v and w share an edge. For each vertex v of G, the set of vertices $\{v : w \sim v\}$ is denoted by ∂v. The spin value ($+1$ or -1) at a vertex v is denoted by σ_v. Now define, for each vertex v and each $\alpha \in \{-1, +1\}^{\partial v}$, the distribution $q_v^\alpha = q_{v;\beta,h}^\alpha$, on $\{-1, +1\}$:

$$\begin{aligned} q_v^\alpha(+1) &= \frac{\exp(\beta(h + \sum_{w\sim v} \alpha_w))}{\exp(\beta(h + \sum_{w\sim v} \alpha_w)) + \exp(-\beta(h + \sum_{w\sim v} \alpha_w))}, \\ q_v^\alpha(-1) &= \frac{\exp(-\beta(h + \sum_{w\sim v} \alpha_w))}{\exp(\beta(h + \sum_{w\sim v} \alpha_w)) + \exp(-\beta(h + \sum_{w\sim v} \alpha_w))}. \end{aligned} \tag{3}$$

Let V denote the set of vertices of G. An Ising distribution on G (with parameters β and h) is a probability distribution $\mu_{\beta,h}$ on $\{-1, +1\}^V$ which satisfies, for each vertex v and each $\eta \in \{-1, +1\}$,

$$\mu_{\beta,h}(\sigma_v = \eta | \sigma_w, w \neq v) = q_v^{\sigma_{\partial v}}(\eta), \qquad \mu_{\beta,h}\text{-a.s.} \tag{4}$$

In this (usual) setup, the spin values are assigned to the vertices. One can define an Ising model with spins assigned to the edges, by replacing G by its cover graph (i.e., the graph whose vertices correspond with the edges of G, and where two vertices share an edge if the edges of G to which these vertices correspond, have a common endpoint).

In the case where G is the d-dimensional cubic lattice $\mathbb{Z}^d$, with $d \geq 2$, it is well known that there is a critical value $\beta_c \in (0, \infty)$ such that the following holds: if $\beta < \beta_c$, there is a unique distribution satisfying (4). If $\beta > \beta_c$ and $h = 0$, there is more than one distribution satisfying (4). A similar result (but with a different value of β_c) holds for the edge version of the model.

Let $b > a > 0$. An $\{a, b\}$*-valued Ising model* is obtained from the usual Ising model by reading a for -1 and b for $+1$. More precisely, if $(\sigma_v, v \in V)$ has an Ising distribution and, for each $v \in V$, $t(v)$ is defined to be a if $\sigma_v = -1$ and b if $\sigma_v = +1$, then we say that $(t(v), v \in V)$ are $\{a, b\}$-valued Ising variables. A similar definition holds for the situation where the spins are assigned to the edges.

A special case of our main result is the following extension of Theorem 1.1 to the Ising model.

THEOREM 1.2. *Let $b > a > 0$ and $d \geq 2$. If $(t(v), v \in \mathbb{Z}^d)$ are $\{a, b\}$-valued Ising variables with inverse temperature $\beta < \beta_c$ and external field h, then there is a constant $C > 0$ such that for all v with $|v| \geq 2$,*

$$\operatorname{Var}(T(0, v)) \leq C \frac{|v|}{\log|v|}. \tag{5}$$

The analog of this result holds for the case where the values a, b are assigned to the edges.

1.2. *Markov random fields with high-noise condition.* Let $(\sigma_v, v \in \mathbb{Z}^d)$, be a translation invariant Markov random field taking values in $W^{\mathbb{Z}^d}$ where W is a finite set. Let $v \in \mathbb{Z}^d$. For each $w \in W$ define (see [7])

$$\gamma_w = \min_{\eta \in W^{\partial v}} \mathbb{P}(\sigma_v = w | \sigma_{\partial v} = \eta).$$

Further, define

$$\gamma = \sum_{w \in W} \gamma_w.$$

Note that the definition of γ_w and γ does not depend on the choice of v. Häggström and Steif [7] studied the existence of finitary codings (and exacts simulations) of Markov random fields under the following high-noise (HN) condition (see also [6] and [21]).

DEFINITION 1.3 (HN condition). A translation invariant Markov random field on $\mathbb{Z}^d$ satisfies the HN condition, if

$$\gamma > \frac{2d - 1}{2d}.$$

We will show that the following theorem is a consequence of our main result.

THEOREM 1.4. *Let $d \geq 2$ and let $(\sigma_v, v \in \mathbb{Z}^d)$ be a translation invariant Markov random field taking finitely many, strictly positive values. If this Markov random field satisfies the HN condition, then, for the first-passage percolation model with $t(v) = \sigma_v$, $v \in \mathbb{Z}^d$, there is a constant $C > 0$ such that for all v with $|v| \geq 2$,*

$$\mathrm{Var}(T(0, v)) \leq C \frac{|v|}{\log|v|}. \tag{6}$$

The analog of this result holds for the edge version of the model.

REMARK. The HN condition for the edge version is a natural modification of that in Definition 1.3. For instance, the $2d$ in the numerator and the denominator of the right-hand side of the inequality in Definition 1.3 is the number of nearest-neighbor vertices of a given vertex, and will be replaced by $4d - 2$ (which is the number of edges sharing an endpoint with a given edge).

1.3. *The minimal number of sign changes in an Ising pattern.* In Section 1.1, the collection of random variables $(t(v), v \in \mathbb{Z}^d)$ itself had an Ising distribution (with -1 and $+1$ translated to a, resp., b). A quite different first-passage percolation process related to the Ising model is the one, studied by Higuchi and Zhang [9], where one counts the minimal number of sign changes from a vertex v to a vertex w in an Ising configuration.

For $\beta < \beta_c$, let $\theta(\beta, h)$ denote the probability that 0 belongs to an infinite $+$ cluster, and let

$$h_c(\beta) = \sup\{h : \theta(\beta, h) = 0\}.$$

For $d = 2$, it was proved in [8] that $h_c(\beta) > 0$.

Using our general results, we will prove (in Section 2) the following extension of Theorem 1.1.

THEOREM 1.5. *Let the collection of random variables $(\sigma_v, v \in \mathbb{Z}^2)$ have an Ising distribution with parameters $\beta < \beta_c$ and external field h, with $|h| < h_c$. Define, for each edge $e = (v_1, v_2)$,*

$$t(e) = \begin{cases} 1, & \text{if } \sigma_{v_1} \neq \sigma_{v_2}, \\ 0, & \text{if } \sigma_{v_1} = \sigma_{v_2}. \end{cases}$$

For the first-passage percolation model with these t-values, there is a $C > 0$ such that for all v with $|v| \geq 2$,

$$\mathrm{Var}(T(0, v)) \leq C \frac{|v|}{\log|v|}. \tag{7}$$

REMARK. Higuchi and Zhang [9] give a concentration result for this model (see Theorem 2 in [9]). Their method is very different from ours. [It is interesting to note that the paragraph below (1.11) in their paper suggests that Talagrand-like inequalities are not applicable to the Ising model.] The upper bound for the variance of $T(0, v)$ which follows from their concentration result is (a "logarithmic-like" factor) *larger* than linear. For earlier results on this and related models, see the Introduction in [9].

1.4. *Statement of the main results*. Our main results, Theorems 1.6 and 1.7, involve t-variables that can be represented by (or "encoded" in terms of) i.i.d. finite-valued random variables in a suitable way, satisfying certain conditions. These conditions are of the same flavor as (but somewhat different from) those in Section 2 in [20].

We first need some notation and terminology. Let S be a finite set, and I a countably infinite set. Let W be a finite subset of I. If $x \in S^I$, we write x_W to denote the tuple $(x_i, i \in W)$. If $h : S^I \to \mathbb{R}$ is a function, and $y \in S^W$, we say that y determines the value of h if $h(x) = h(x')$ for all x, x' satisfying $x_W = x'_W = y$.

Let $X_i, i \in I$, be i.i.d. S-valued random variables. We say that the random variables $t(v), v \in \mathbb{Z}^d$, are represented by the collection $(X_i, i \in I)$, if, for each $v \in \mathbb{Z}^d$, $t(v)$ is a function of $(X_i, i \in I)$. The formulation of our main theorems involve certain conditions on such a representation:

- *Condition* (i): There exist $c_0 > 0$ and $\varepsilon_0 > 0$ such that for each $v \in \mathbb{Z}^d$ there is a sequence $i_1(v), i_2(v), \dots$ of elements of I, such that for all $k = 1, 2, \dots,$

$$P\big((X_{i_1(v)}, \dots, X_{i_k(v)}) \text{ does not determine } t(v)\big) \leq \frac{c_0}{k^{3d+\varepsilon_0}}. \tag{8}$$

- *Condition* (ii):

$$\begin{gathered} \exists \alpha > 0 \ \forall v, w \in \mathbb{Z}^d \ \forall k < \alpha |v - w| \\ \{i_1(v), \dots, i_k(v)\} \cap \{i_1(w), \dots, i_k(w)\} = \varnothing. \end{gathered} \tag{9}$$

- *Condition* (iii): The distribution of the family of random variables $(t(v), v \in \mathbb{Z}^d)$ is translation-invariant.

We say that the family of random variables $(t(v), v \in \mathbb{Z}^d)$ has a representation satisfying conditions (i)–(iii), if there are S, I and i.i.d. S-valued random variables $X_i, i \in I$ as above, such that the t-variables are functions of the X-variables satisfying conditions (i)–(iii) above.

Analogs of these definitions for t-variables indexed by the edges of $\mathbb{Z}^d$ can be given in a straightforward way.

Now we are ready to state our main theorem.

THEOREM 1.6. *Let $b > a > 0$, and let, with $d \geq 2$, $(t(v), v \in \mathbb{Z}^d)$ be a family of random variables that take values in the interval $[a, b]$ and have a representation* *satisfying conditions* (i)–(iii) *above*. *Then there is a* $C > 0$, *such that for all* $v \in \mathbb{Z}^d$ *with* $|v| \geq 2$,

$$\mathrm{Var}(T(0, v)) \leq \frac{C|v|}{\log|v|}. \tag{10}$$

The analog for the bond version of this result also holds.

If the t-variables can take values equal or arbitrarily close to 0, we need a stronger version of condition (i) and extra condition (iv) (see below).

By an optimal path from v to w, we mean a path π from v to w such that $T(\pi) \leq T(\pi')$ for all paths π' from v to w.

- *Condition* (i′): There exist $c_0 > 0$, $\varepsilon_0 > 0$ and $\varepsilon_1 > 0$, such that for each $v \in \mathbb{Z}^d$ there is a sequence $i_1(v), i_2(v), \dots$ of elements of I, such that for all $k = 1, 2, \dots,$

$$P\big((X_{i_1(v)}, \dots, X_{i_k(v)}) \text{ does not determine } t(v)\big) \leq c_0 \exp(-\varepsilon_0 k^{\varepsilon_1}). \tag{11}$$

- *Condition* (iv): There exist $c_1, c_2, c_3 > 0$ such that for all vertices v, w the probability that there is no optimal path π from v to w with $|\pi| \leq c_1|v - w|$ is at most $c_2 \exp(-c_3|v - w|)$.

THEOREM 1.7. *Let* $b > 0$, *and let*, *with* $d \geq 2$, $(t(v), v \in \mathbb{Z}^d)$ *be a collection of random variables taking values in the interval* $[0, b]$, *and having a representation satisfying conditions* (i′), (ii), (iii) *and* (iv) *above*. *Then there is a* $C > 0$, *such that for all* $v \in \mathbb{Z}^d$ *with* $|v| \geq 2$,

$$\mathrm{Var}(T(0, v) \leq \frac{C|v|}{\log|v|}. \tag{12}$$

The analog of this result for the bond version of the model also holds.

REMARKS.

(a) Note that condition (iii) is in terms of the t-variables only: we do *not* assume that the index set I has a "geometric" structure and that the t-variables are "computed" from the X-variables in a "translation-invariant" way with respect to that structure (and the structure of $\mathbb{Z}^d$).

(b) The goal of our paper is to show that the main result in [2], although its proof heavily uses inequalities concerning independent random variables, can be extended to an interesting class of dependent first-passage percolation models. In the setup of the above conditions (i), (ii), (iii), (i′) and (iv), we have aimed to obtain fairly general Theorems 1.6 and 1.7, without becoming too general (which would give rise to so many extra technicalities that the main line of argument would be obscured). For instance, from the proofs it will be clear that there is a kind of "trade-off" between conditions (i) and (ii): one may simultaneously strengthen the first and weaken the second condition.

Also, if the bound in condition (i′) is replaced by a polynomial bound with sufficiently high degree, Theorem 1.7 would still hold (but more explanation would be needed in Section 5). Since the main motivation for adding this theorem to Theorem 1.6 is to handle the interesting Ising sign-change model studied by Higuchi and Zhang [for which we know that condition (i′) holds] we have not replaced condition (i′) by a weaker condition.

2. Proofs of Theorems 1.2, 1.4 and 1.5 from Theorems 1.6 and 1.7.

2.1. *Proof of Theorem* 1.2. In [20], the notion "nice finitary representation" has been introduced in the context of two-dimensional random fields. See conditions (i)–(iv) in Section 2 of that paper. In Section 2 (see in particular Theorem 2.3 in that paper), it is shown that the Ising model with $\beta < \beta_c$ has such a representation. (See also [21].) The key ideas and ingredients are exact simulation by coupling from the past (see [17] and [21]), and a well-known result by Martinelli and Olivieri [16] that under a natural dynamics (single-site updates; Gibbs sampler) the system has exponential convergence to the Ising distribution. The random variables used to execute these updates are taken as the X-variables in the definition of a representation.

Condition (ii) in [20] is somewhat weaker than our current condition (i). However, as shown in [20] (see the arguments between Theorems 2.3 and 2.4 in [20]), the above mentioned exponential convergence shows that the Ising model satisfies an even stronger bound, namely condition (i′) in our paper.

Condition (iii) in [20] corresponds with our condition (ii), and condition (iv) in [20] is stronger than our condition (iii).

In [20], only the two-dimensional case is treated (because the applications are to percolation models where typical two-dimensional methods are used) but its arguments concerning "nice finitary representations" for the Ising model extend immediately to higher dimensions.

From the above considerations, it follows that the Ising models in the statement of our Theorem 1.2 indeed have a representation satisfying our conditions (i)–(iii). Application of Theorem 1.6 now gives Theorem 1.2.

2.2. *Proof of Theorem* 1.4. The argument is very similar to that in the proof of Theorem 1.2. Therefore, we only mention the points that need extra attention.

As in the proof of Theorem 1.2, the role of the X-variables in Section 1.4 is played by the i.i.d. random variables driving a single-site update scheme (Gibbs sampler). In Theorem 1.2, a form of exponential convergence for the Gibbs sampler was used. This exponential convergence came from a result in [16]. In the current situation, the exponential convergence is, as shown in Proposition 2.1 in [7], a consequence of the HN condition. This exponential convergence implies (again, as in the case of Theorem 1.2) condition (i) [and, in fact, the stronger condition (i′)] in Section 1.4. Condition (iii) is obvious, and condition (ii) follows easily (as in the proof of Theorem 1.2) from the general setup of the Gibbs sampler. So, again, we now apply Theorem 1.6 to obtain Theorem 1.4.

2.3. *Proof of Theorem* 1.5. Since $\beta < \beta_c$, the collection $(\sigma_v, v \in \mathbb{Z}^2)$, has (as pointed out in the proof of Theorem 1.2) a representation satisfying conditions (i), (ii) and (iii). In fact, as noted in the proof of Theorem 1.2, it even satisfies the stronger form (i′) of condition (i). Since $t(e)$ is a function of the σ-values of the two endpoints of e, it follows immediately that the collection $(t(e), e \in E)$ (where E denotes the set of edges of the lattice $\mathbb{Z}^2$) satisfies the (bond analog of) the conditions (i′), (ii) and (iii). The fact that (iv) is satisfied follows immediately from Lemma 6 [and (1.9)] in [9]. Theorem 1.5 now follows from (the bond version of) Theorem 1.7.

3. Ingredients for the proof of Theorem 1.6.

3.1. *An inequality by Talagrand.* Let S be a finite set and n a positive integer. Assign probabilities p_s, $s \in S$, to the elements of S. Let μ be the corresponding product measure on $\Omega := S^n$.

Let f be a function on Ω, and let $\|f\|_1$ and $\|f\|_2$ denote the L_1-norm and L_2-norm of f w.r.t. the measure μ:

$$\|f\|_1 := \sum_{x\in\Omega} \mu(x)|f(x)|;$$

$$\|f\|_2 := \sqrt{\sum_{x\in\Omega} \mu(x)|f(x)|^2}.$$

The notation $\bar{f}_i$ is used for the conditional expectation of f given all coordinates except the ith. More precisely, for $x = (x_1, \ldots, x_n) \in S^n$ we define

$$\bar{f}_i(x) := \sum_{s\in S} p_s f(x_1, \ldots, x_{i-1}, s, x_{i+1}, \ldots, x_n).$$

Further, we define the function $\Delta_i f$ on Ω by

$$(\Delta_i f)(x) = f(x) - \bar{f}_i(x), \qquad x \in \Omega. \tag{13}$$

NOTATIONAL REMARK. Often we work with the alternative, equivalent, description that we have n independent random variables, say $Z_1, \ldots, Z_n$, with $P(Z_i = s) = p_s$, $s \in S, 1 \leq i \leq n$. To emphasize the identity of the random variables involved, we then often use the notation Δ_{Z_i} instead of Δ_i.

A key ingredient in [2] and in our paper is the following inequality for the case $|S| = 2$ by Talagrand, a far-reaching extension of an inequality by Kahn, Kalai and Linial [12].

THEOREM 3.1 (Talagrand [19], Theorem 1.5). *There is a constant $K > 0$ such that for each n and each function f on $\{0, 1\}^n$,*

$$\mathrm{Var}(f) \le K \log\left(\frac{2}{p(1-p)}\right) \sum_{i=1}^{n} \frac{\|\Delta_i f\|_2^2}{\log(e\|\Delta_i f\|_2/\|\Delta_i f\|_1)}, \tag{14}$$

where (*in the notation in the beginning of this section*) $p = p_1 = 1 - p_0$, *and where* $\mathrm{Var}(f)$ *denotes the variance of f w.r.t. the measure μ.*

In the literature, (partial) extensions of this inequality and inequalities of related flavor, to the case $|S| > 2$ have been given; see, for example, [18] and [1]. The following theorem (see [14]) states the most "literal" extension of Theorem 3.1 to the case $|S| > 2$. (In [14], an extended version of Beckner's inequality, a key ingredient in the proof of Theorem 3.1, is used, and the proof of Talagrand is followed, with appropriate adaptations, to obtain the extension of Theorem 3.1.) To make comparison of our line of arguments with that in [2] as clear as possible, it is this extension we will use. (Moreover, if instead of Theorem 3.2 we would use the modified Poincaré inequalities in [1], this would not simplify our proof of Theorem 1.6.)

THEOREM 3.2 ([14], Theorem 1.3). *There is a constant $K > 0$ such that for each finite set S, each $n \in \mathbb{N}$ and each function f on S^n the following holds*:

$$\mathrm{Var}(f) \le K \log\left(\frac{1}{\min_{s\in S} p_s}\right) \sum_{i=1}^{n} \frac{\|\Delta_i f\|_2^2}{\log(e\|\Delta_i f\|_2/\|\Delta_i f\|_1)}. \tag{15}$$

3.2. *Greedy lattice animals.* The subject of this subsection played no role in the treatment of the first-passage percolation model with independent t-variables in [2], but turns out to be important in our treatment of dependent t-variables.

Consider, for $d \ge 2$, the d-dimensional cubic lattice. A lattice animal (abbreviated as l.a.) is a finite connected subset of $\mathbb{Z}^d$ containing the origin. Let X_v, $v \in \mathbb{Z}^d$, be i.i.d. nonnegative random variables with common distribution F. Define

$$N(n) := \max_{\zeta\,:\,\zeta \text{ l.a. with } |\zeta|=n} \sum_{v\in\zeta} X_v,$$

where the maximum is over all lattice animals of size n.

The subject was introduced by Cox et al. [3]. The asymptotic behavior, as $n \to \infty$ of $N(n)$ has been studied in that and several other papers (see, e.g., [4] and [10]). For our purpose, the following result by Martin [15] is very suitable.

THEOREM 3.3 (Martin [15], Theorem 2.3). *There is a constant C such that for all n and for all F that satisfy*

$$\int_0^\infty (1 - F(x))^{1/d}\, dx < \infty,$$

$$E\left(\frac{N(n)}{n}\right) \le C \int_0^\infty (1 - F(x))^{1/d}\, dx. \tag{16}$$

Martin [15] says considerably more than this, but the above is sufficient for our purpose.

3.3. *A randomization tool.* As in [2] we need, for technical reasons, a certain "averaging" argument: extra randomness is added to the system to make it more tractable. To handle this extra randomness appropriately, the following lemma from [2] is used.

LEMMA 3.4 (Benjamini, Kalai and Schramm [2], Lemma 3). *There is a constant $c > 0$ such that for every $m \in \mathbb{N}$ there is a function*

$$g = g_m : \{0,1\}^{m^2} \to \{0, 1, \dots, m\},$$

which satisfies properties (i) *and* (ii) *below*:

(i) *For all $i = 1, \dots, m^2$ and all $x \in \{0,1\}^{m^2}$*,

$$\left|g_m(x^{(i)}) - g_m(x)\right| \le 1, \tag{17}$$

where $x^{(i)}$ denotes the element of $\{0,1\}^{m^2}$ that differs from x only in the ith coordinate.

(ii)

$$\max_k \mathbb{P}(g(y) = k) \le c/m, \tag{18}$$

where y is a random variable uniformly distributed on $\{0,1\}^{m^2}$.

4. Proof of Theorem 1.6. To keep our formulas compact, we will use constants $C_1, C_2, \dots$. The precise values of these constants do not matter for our purposes. Some of them depend on a, b, the dimension d, the distribution of the X-variables (in terms of which the t-variables are represented), or the constants in the conditions (i), (i′), (ii), (iii) and (iv) in Section 1.4. However, they do not (and obviously should not) depend on the choice of v in the statement of the theorem.

4.1. *Informal sketch.* The detailed proof is given in the next subsection. Now we first give a very brief and rough summary of the proof of the main result in [2] (listed as Theorem 1.1 in our paper), and then informally (and again briefly) indicate the extra problems that arise in our situation where the t-variables are dependent.

Let γ be the path from 0 to v for which the sum of the t-variables is minimal. (If more than one such path exists, choose one of these by a deterministic procedure.)

Since the value of each t-variable is at least $a > 0$ and at most b, it is clear that the number of edges of γ is at most a constant c times $|v|$.

In [2] the t-variables are independent, and Talagrand's inequality (Theorem 3.1) is applied with $f = T(0, v)$ and with each i denoting an edge e. From the definitions, it is clear that $\Delta_i f$ is roughly the change of $T(0, v)$ caused by changing $t(e)$. Moreover, a change of $t(e)$ can only cause a change of $T(0, v)$ if, before or after the change, e is on the above mentioned path γ. So, ignoring the denominator in Talagrand's inequality, one gets the (linear) bound

$$\operatorname{Var}(T(0, v)) \le C_1 \mathbb{E}\left[\sum_{e \in \gamma} (b - a)^2\right] \le c(b - a)^2 |v|. \tag{19}$$

It turns out that, by introducing additional randomness in an appropriate way, without changing the variance (see Lemma 3.4), the $\|\Delta_i f\|_2 / \|\Delta_i f\|_1$ in the denominator in the right-hand side of Talgrand's inequality becomes (uniformly in i) larger than $|v|^\beta$ for some $\beta > 0$, thus giving the $\log|v|$ (and hence, the sublinearity) in Theorem 1.1.

In our situation, the underlying independent random variables are the X_i, $i \in I$ (by which the *dependent* t-variables are represented). Application of Talagrand-type inequalities to these variables has the complication that changing one X-variable changes a (random) set of possibly many t-variables. Taking the square of the effect complicates this further. Nevertheless, it turns out that by suitable decompositions of the summations, and by block arguments (rescaling), one finally gets, instead of (19) a bound in terms of ("rescaled") greedy lattice animals which, by the result of Martin in Section 3, is still linear in $|v|$.

To handle the denominator in the Talagrand-type inequality, we use additional randomness, as in [2]. Again, the fact that changing an X-variable can have effect on many t-variables complicates the analysis, but this complication is easier to handle than that for the numerator mentioned above.

4.2. *Detailed proof.* We give the proof for the site version of Theorem 1.6. The proof for the bond version is obtained from it by a straightforward, step-by-step translation.

NOTATIONAL REMARK. The cardinality of a set V will be indicated by $|V|$.

We start by stating a simple but important observation (a version of which was also used in [2]). A finite path π is called an optimal path, or a geodesic, if there is no path $\pi' \neq \pi$ with the same starting and endpoint as π, for which $T(\pi') < T(\pi)$.

OBSERVATION 4.1. *Since the t-variables are bounded away from 0 and ∞, there is a constant $C_2 > 0$ such that for every positive integer n and every $w \in \mathbb{Z}^d$ the following hold*:

(a) *Each geodesic has at most $C_2 n$ vertices in the box $w + [-n, n]^d$.*

(b) *Each geodesic which starts at* 0 *and ends at* w *has at most* $C_2|w|$ *vertices.*

Let $X_i, i \in I$, be the independent random variables in terms of which the variables $(t(v), v \in \mathbb{Z}^d)$ are represented. So $T(0, v)$ is a function of the X-variables. As we said in the informal sketch, we introduce extra randomness, in the same way as in [2]: fix $m := \lfloor |v|^{1/4} \rfloor$. Let $(y_i^j, i = 1, \dots, m^2, j = 1, \dots, d)$ be a family of independent random variables, each taking value 0 or 1 with probability $1/2$. The family of y_i^j's is also taken independently of the X-variables. Define, for $j = 1, \dots, d$,

$$y^j = (y_1^j, \dots, y_{m^2}^j).$$

Each y^j is uniformly distributed on $\{0, 1\}^{m^2}$, and will play the role of the y in Lemma 3.4. We simply write Y for the collection $(y_i^j, i = 1, \dots, m^2, j = 1, \dots, d)$ and X for the collection $(X_i, i \in I)$.

Let

$$z(Y) = (g(y^1), \dots, g(y^d)) \tag{20}$$

with $g = g_m$ as in Lemma 3.4.

To shorten notation, we will write f for $T(O, v)$ and $\tilde{f}$ for the passage time between the vertices that are obtained from 0 and v by a (random) shift over the vector $z(Y)$:

$$\tilde{f} = T\big(z(Y), v + z(Y)\big). \tag{21}$$

Note that f is completely determined by X, while $\tilde{f}$ depends on X as well as Y.

By translation invariance [see condition (iii)], for every $w \in \mathbb{Z}^d$, $T(0, v)$ has the same distribution as $T(w, v + w)$. Hence, by conditioning on Y and using that Y is independent of the t-variables, it follows that $\tilde{f}$ has the same distribution as f. In particular,

$$\mathrm{Var}(f) = \mathrm{Var}(\tilde{f}). \tag{22}$$

Theorem 3.2 gives (see the Remarks below)

$$\begin{aligned} \mathrm{Var}(\tilde{f}) \le C_3 & \sum_{i=1,\dots,m^2, j=1,\dots,d} \frac{\|\Delta_{y_i^j} \tilde{f}\|_2^2}{1 + \log(\|\Delta_{y_i^j} \tilde{f}\|_2 / \|\Delta_{y_i^j} \tilde{f}\|_1)} \\ & + C_3 \frac{\sum_{i \in I} \|\Delta_{X_i} \tilde{f}\|_2^2}{1 + \min_{i \in I} \log(\|\Delta_{X_i} \tilde{f}\|_2 / \|\Delta_{X_i} \tilde{f}\|_1)}. \end{aligned} \tag{23}$$

REMARKS.

(a) At first sight, Theorem 3.2 is not applicable in the current situation where we have two types of random variables: X_i's and y_i^j's. However, by a straightforward argument, "pairing" each variable y_i^j, $i = 1, \dots, m^2$, $j = 1, \dots, d$, with an independent "dummy" variable X_i^j (with the same distribution as the "ordinary" X-variables), and each variable X_i, $i \in I$, with an independent "dummy" variable y_i (with the same distribution as the "ordinary" y-variables), it is easy to see that Theorem 3.2 is indeed applicable here.

(b) Note that the statement of Theorem 3.2 is formulated for finite n. Combined with a standard limit argument, it gives (23).

We will handle, in separate subsections, the first term of (23), the numerator of the second term, and the denominator of the second term.

4.2.1. *The first term in* (23). By (20), (21) and (17) it follows that $|\Delta_{y_i^j} \tilde{f}|$ is at most a constant C_4, so that we have the following lemma.

LEMMA 4.2. *The first term in* (23) *is at most*

$$\le dC_4 m^2 = dC_4 |v|^{1/2}. \tag{24}$$

4.2.2. *The denominator of the second term in* (23). In this subsection we write, for notational convenience, $\Delta_i \tilde{f}$ for $\Delta_{X_i} \tilde{f}$, where $i \in I$.

If $w, w' \in \mathbb{Z}^d$ we write $\gamma_{w,w'}$ for the path π minimizing $\sum_{w \in \pi} t(w)$. If there is more than one such path, we use a deterministic, translation-invariant way to select one. If $w = 0$ and w' is our "fixed" v, we write simply γ for $\gamma_{0,v}$.

Recall that $z = z(Y)$ is the random shift. We write $\gamma(z)$ for $\gamma_{z,v+z}$.

Also recall the definitions and notation in Section 1.4. If $w \in \mathbb{Z}^d$ and $j \in I$, we say that w needs j if $j = i_k(w)$ for some positive integer k, and $X_{i_1(w)}, \dots, X_{i_{k-1}(w)}$ does not determine $t(w)$.

By a well-known second-moment argument we have, for each $j \in I$,

$$\frac{\|\Delta_j \tilde{f}\|_2}{\|\Delta_j \tilde{f}\|_1} \ge \frac{1}{\sqrt{\mathbb{P}(\Delta_j \tilde{f} \neq 0)}}. \tag{25}$$

Note that, given $z(Y)$ and all $X_i, i \in I \setminus \{j\}$, there is a, possibly nonunique, $s = s(j, X, Y) \in S$ such that $\tilde{f}$ (now considered as a function of X_j only) takes its smallest value at $X_j = s$. Further note that if $\Delta_j \tilde{f} \neq 0$ then, after replacing the value of X_j by s, we have $\Delta_j \tilde{f} < 0$. So we get

$$\mathbb{P}(\Delta_j \tilde{f} < 0) \ge \mathbb{P}(\Delta_j \tilde{f} \neq 0) \min_{r \in S} \mathbb{P}(X_j = r),$$

and hence

$$\mathbb{P}(\Delta_j \tilde{f} \neq 0) \le \frac{\mathbb{P}(\Delta_j \tilde{f} < 0)}{\min_{r \in S} \mathbb{P}(X_j = r)}. \tag{26}$$

Moreover, it follows from the definitions that if $\Delta_j \tilde{f} < 0$, there is a w on $\gamma(z)$ such that a certain change of X_j causes a change of $t(w)$. By this and (26), we have

$$\begin{aligned} \mathbb{P}(\Delta_j \tilde{f} \neq 0) &\leq C_5 \sum_{w \in \mathbb{Z}^d} \mathbb{P}\big(w \in \gamma(z), w \text{ needs } j\big) \\ &\leq C_5 \sum_{w \in \mathbb{Z}^d} \min\big(\mathbb{P}\big(w \in \gamma(z)\big), \mathbb{P}(w \text{ needs } j)\big). \end{aligned} \tag{27}$$

Recall the definition of m in the paragraph following Observation 4.1. Let $w \in \mathbb{Z}^d$ and consider the box $B_m(w) := w + [-m, m]^d$. We have

$$\mathbb{P}\big(w \in \gamma(z)\big) = \mathbb{P}(w - z \in \gamma).$$

By the construction of z, and (17), $w - z$ takes values in the above mentioned box $B_m(w)$. Also by the construction of z, and (18), each vertex of the box has probability $\leq C_6/m^d$ to be equal to $w - z$. Moreover, by Observation 4.1 at most $C_7 m$ vertices in the box are on γ. Hence, since γ is independent of z, it follows (by conditioning on γ) that

$$\mathbb{P}\big(w \in \gamma(z)\big) \leq C_7 m \frac{C_6}{m^d} \leq C_8 |v|^{-(d-1)/4}. \tag{28}$$

Further, by condition (i), we have

$$\mathbb{P}(w \text{ needs } j) \leq \frac{c_0}{r_w(j)^{3d+\varepsilon_0}}, \tag{29}$$

where $r_w(j)$ (which we call the rank of j) is the positive integer k for which $i_k(w) = j$.

By (27), (28) and (29), we have, for every K,

$$\mathbb{P}(\Delta_j \tilde{f} \neq 0) \leq C_9 \left(|v|^{-(d-1)/4} |\{w : r_w(j) < K\}| + \sum_{k=K}^{\infty} \frac{|\{w : r_w(j) = k\}|}{k^{3d+\varepsilon_0}} \right). \tag{30}$$

Now, condition (ii) implies, for each $j \in I$ and each $k > 0$,

$$|\{w : r_w(j) < k\}| \leq C_{10} k^d. \tag{31}$$

Hence, the first term between the brackets in (30) is at most

$$C_{10} |v|^{-(d-1)/4} K^d. \tag{32}$$

Further, using again (31) (and summation by parts) the sum over k in (30) is at most

$$C_{11} \sum_{k=K}^{\infty} \frac{k^d}{k^{3d+\varepsilon_0+1}} \leq C_{12} K^{-2d-\varepsilon_0}. \tag{33}$$

Combining (30), (32) and (33), we get

$$\mathbb{P}(\Delta_j \tilde{f} \neq 0) \leq C_{13}\big(|v|^{-(d-1)/4} K^d + K^{-2d-\varepsilon_0}\big). \tag{34}$$

Now take for K the smallest positive integer satisfying $K^d \geq |v|^{(d-1)/8}$ and insert this in (34). This gives

$$\mathbb{P}(\Delta_j \tilde{f} \neq 0) \leq C_{14} |v|^{-(d-1)/8}, \tag{35}$$

which together with (25) yields the following lemma.

LEMMA 4.3. *There is a constant $C_{15} > 0$ such that for all $v \in \mathbb{Z}^d$ the denominator of the second term in* (23) *is larger than or equal to*

$$C_{15} \log|v|.$$

4.2.3. *The numerator of the second term in* (23), *and completion of the proof of Theorem* 1.6. As in the previous subsection, we write Δ_j for Δ_{X_j}, where $j \in I$.

By the definition of $\tilde{f}$ (and of the norm $\|\cdot\|_2$), we rewrite

$$\sum_{j \in I} \|\Delta_j \tilde{f}\|_2^2 = \sum_{j \in I} \mathbb{E}\big[\big(\Delta_j T\big(z(Y), z(Y) + v\big)\big)^2\big]. \tag{36}$$

By taking the expectation outside the summation, conditioning on Y (and using that Y is independent of the t-variables) and then taking the expectation back inside the summation, it is clear that the right-hand side of (36) is smaller than or equal to

$$\max_{x \in \mathbb{Z}^d} \sum_{j \in I} \mathbb{E}\big(\big(\Delta_j T(x, x+v)\big)^2\big). \tag{37}$$

We will give an upper bound for the sum in (37) for the case $x = 0$. From the computations, it will be clear that this upper bound does not use the specific choice of x, and hence holds for all x.

In the case $x = 0$, the sum in (37) is, by definition, of course

$$\sum_{j \in I} \|\Delta_j f\|_2^2. \tag{38}$$

Let X'_j be an auxiliary random variable that is independent of the X-variables and has the same distribution. Let X denote the collection of random variables $(X_i, i \in I)$, and X' the collection obtained from the collection X by replacing X_j by X'_j. By the definition of $\Delta_j f$ (and standard arguments), we have

$$\begin{aligned} \mathbb{E}_j((\Delta_j f)^2) &= \tfrac{1}{2}\mathbb{E}_{j,j'}\big[\big(f(X) - f(X')\big)^2\big] \\ &= \mathbb{E}_{j,j'}\big[\big(f(X) - f(X')\big)^2 I\big(f(X) < f(X')\big)\big], \end{aligned} \tag{39}$$

where $\mathbb{E}_j$ denotes the expectation with respect to X_j, and $\mathbb{E}_{j,j'}$ denotes the expectation with respect to X_j and X'_j. [So, (39) is a function of the collection $(X_i, i \in I, i \neq j)$.]

Let γ be the optimal path, as defined in the beginning of Section 4.2.2, w.r.t. the t-variables corresponding with the family X. Let w be a vertex. Observe that a change of $t(w)$ does not increase f if w is not on γ, and increases f by at most $b-a$ if w is on γ. By this observation, and a similar argument as used for (27), we have

$$(f(X') - f(X))I(f(X) < f(X')) \leq (b-a) \sum_{w \in \gamma} I(w \text{ needs } j), \tag{40}$$

and hence

$$\begin{aligned}
&(f(X) - f(X'))^2 I(f(X) < f(X')) \\
&\qquad \leq (b-a)^2 \sum_{u,w \in \gamma} I(u \text{ needs } j, w \text{ needs } j).
\end{aligned} \tag{41}$$

Since $\|\Delta_j f\|_2^2$ is the expectation w.r.t. the X_i, $i \neq j$, of $\mathbb{E}_j((\Delta_j f)^2)$, we have, by (39) and (41), that

$$\|\Delta_j f\|_2^2 \leq (b-a)^2 \mathbb{E}\Bigg[\sum_{u,w \in \gamma} I(u \text{ needs } j, w \text{ needs } j)\Bigg]. \tag{42}$$

To bound the right-hand side of (42), recall the definition [below (29)] of $r_w(j)$ (with $j \in I$ and $w \in \mathbb{Z}^d$), and note that, by condition (i) in Section 1.4, we have, on an event of probability 1,

$$\begin{aligned}
&\sum_{u,w \in \gamma} I(u \text{ and } w \text{ need } j) \\
&\qquad = \sum_{k=1}^{\infty} \sum_{u,w \in \gamma} I\big(u \text{ and } w \text{ need } j, \max(r_u(j), r_w(j)) = k\big) \\
&\qquad \leq 2 \sum_{k=1}^{\infty} \sum_{u \in \gamma} \sum_{w \in \gamma} I\big(u \text{ and } w \text{ need } j, r_u(j) = k, r_w(j) \leq k\big) \\
&\qquad \leq 2 \sum_{k=1}^{\infty} \sum_{u \in \gamma} I\big(u \text{ needs } j, r_u(j) = k\big) |\{w \in \gamma : r_w(j) \leq k\}|.
\end{aligned} \tag{43}$$

By condition (ii), each of the vertices w in the last line of (43) is located in a hypercube of length $C_{16}k$ centered at u. By this and Observation 4.1, it follows that the number of w's in the last line of (43) is at most $C_{17}k$. So we have, with $C_{18} = 2C_{17}$,

$$\sum_{u,w \in \gamma} I(u \text{ and } w \text{ need } j) \leq C_{18} \sum_{k=1}^{\infty} k \sum_{u \in \gamma} I\big(u \text{ needs } j, r_u(j) = k\big),$$

which, together with (42) [and using the definition of $i_k(u)$] gives, after summing over j,

$$\begin{aligned}\sum_{j\in I}\|\Delta_j f\|_2^2 &\le C_{19}\sum_{k=1}^{\infty}k\mathbb{E}\Bigl[\sum_{u\in\gamma}I\bigl(u \text{ needs } i_k(u)\bigr)\Bigr]\\ &= C_{19}\sum_{k=1}^{|v|}k\mathbb{E}\Bigl[\sum_{u\in\gamma}I\bigl(u \text{ needs } i_k(u)\bigr)\Bigr] \\ &\quad + C_{19}\sum_{k>|v|}k\mathbb{E}\Bigl[\sum_{u\in\gamma}I\bigl(u \text{ needs } i_k(u)\bigr)\Bigr].\end{aligned}\tag{44}$$

The sum over $k > |v|$ in the right-hand side of (44) can be bounded very easily as follows: by Observation 4.1(b), all vertices of γ are inside the box $[-C_2|v|, C_2|v|]^d$. Hence, the above-mentioned sum over $k > |v|$ is at most

$$C_{19}\sum_{k>|v|}k\sum_{u\in[-C_2|v|,C_2|v|]^d}\mathbb{P}\bigl(u \text{ needs } i_k(u)\bigr).$$

By condition (i), and since the number of vertices u in this last expression is, of course, of order $|v|^d$, this expression is smaller than or equal to a constant times

$$|v|^d\sum_{k>|v|}k^{1-3d-\varepsilon_0},$$

which is smaller than a constant C_{20}.

To bound the sum over $k \le |v|$ in the right-hand side of (44), observe that, by condition (ii), if a set $V \subset \mathbb{Z}^d$ is such that $|u-u'| \ge C_{21}k$ for all $u, u' \in V$ with $u \ne u'$, then the collection of random variables

$$\bigl(I\bigl(u \text{ needs } i_k(u)\bigr), u \in V\bigr)$$

is independent. With this in mind, we partition, for each k, $\mathbb{Z}^d$ in boxes

$$B_k(w) := [-\lceil C_{21}k\rceil, \lceil C_{21}k\rceil)^d + 2\lceil C_{21}k\rceil w, \qquad w \in \mathbb{Z}^d.$$

We will say that two boxes $B_k(w)$ and $B_k(u)$ are neighbors [where $u = (u_1, \ldots, u_d)$ and $w = (w_1, \ldots, w_d)$] if $\max_{1\le i\le d}|w_i - u_i| = 1$.

By Observation 4.1(a), γ has at most $C_{22}k$ vertices in each of these boxes. Hence, the the sum over $k \le |v|$ in the right-hand side of (44) is at most

$$C_{23}\sum_{k=1}^{|v|}k^2\mathbb{E}\Bigl[\sum_{w:(*)}I\bigl(\exists u \in B_k(w) \text{ s.t. } u \text{ needs } i_k(u)\bigr)\Bigr],\tag{45}$$

where $(*)$ indicates that we sum over all $w \in \mathbb{Z}^d$ with the property that γ has a vertex in $B_k(w)$ or in a neighbor of $B_k(w)$.

Next, partition $\mathbb{Z}^d$ in 2^d classes, as follows:

$$\mathbb{Z}_z := z + 2\mathbb{Z}^d, \qquad z \in \{0,1\}^d.$$

So (45) can be written as

$$C_{23} \sum_{k=1}^{|v|} k^2 \sum_{z \in \{0,1\}^d} \mathbb{E}\Big[\sum_{w\,:\,(**)} I\big(\exists u \in B_k(z+2w) \text{ s.t. } u \text{ needs } i_k(u)\big)\Big], \tag{46}$$

where $(**)$ indicates that we sum over all $w \in \mathbb{Z}^d$ with the property that γ has a point in $B_k(z+2w)$ or in a neighbor of $B_k(z+2w)$.

Now, for each $z \in \{0,1\}^d$, the set

$$\{w \in \mathbb{Z}^d : \gamma \text{ has a point in } B_k(z+2w) \text{ or a neighbor of } B_k(z+2w)\}$$

is a lattice animal and has, for $k \le |v|$, by Observation 4.1(b), at most $C_{24}|v|/k$ elements.

So, from (44)–(46) we get

$$\begin{aligned}\sum_{j\in I} \|\Delta_j f\|_2^2 \le C_{23} \sum_{k=1}^{|v|} k^2 \sum_{z\in\{0,1\}^d} \mathbb{E}\Big[\max_{\mathcal{L}\,:\,|\mathcal{L}|\le C_{24}|v|/k} \sum_{w\in\mathcal{L}} I\big(\exists u \in B_k(z+2w) \\ \text{s.t. } u \text{ needs } i_k(u)\big)\Big] \\ + C_{20},\end{aligned} \tag{47}$$

where the maximum is over all lattice animals $\mathcal{L}$ with size $\le C_{24}|v|/k$.

Now for each z we have, by the observation below (44), that

$$\big(I\big(\exists u \in B_k(z+2w) \text{ s.t. } u \text{ needs } i_k(u)\big), w \in \mathbb{Z}^d\big)$$

is a collection of independent 0–1-valued random variables. For each w, this random variable is 1 with probability less than or equal to

$$|B_k(z+2w)| \max_{u\in\mathbb{Z}^d} \mathbb{P}\big(u \text{ needs } i_k(u)\big) \le \frac{C_{25}k^d}{k^{3d+\varepsilon_0}}, \tag{48}$$

where we used condition (i).

By (47), (48) and Theorem 3.3, we get

$$\begin{aligned}\sum_{j\in I} \|\Delta_j f\|_2^2 &\le C_{20} + C_{26} \sum_{k=1}^{|v|} k^2 \frac{|v|}{k} \bigg(\frac{k^d}{k^{3d+\varepsilon_0}}\bigg)^{1/d} \\ &\le C_{20} + C_{26}|v| \sum_{k=1}^{\infty} k^2 k^{-(3d+\varepsilon_0)/d} \\ &\le C_{27}|v|.\end{aligned} \tag{49}$$

Together with (36)–(38), this gives the following lemma.

LEMMA 4.4. *The numerator of the second term in* (23) *is at most* $C_{27}|v|$.

Lemma 4.4, together with (22), (23), Lemma 4.2 and Lemma 4.3, completes the proof of Theorem 1.6.

5. Proof of Theorem 1.7. The proof is very similar to that of Theorem 1.6 and we only discuss those steps that need adaptation.

First, we define, for $u, w \in \mathbb{Z}^d$, the following modification of $T(u, w)$:

$$\hat{T}(u, w) := \min_{\pi : u \to w, |\pi| \le c_1 |u-w|} T(\pi), \tag{50}$$

where $|\pi|$ is the number of vertices of π, and with c_1 as in condition (iv).

From this definition, it is obvious that

$$\begin{aligned} &|\hat{T}(0, v) - T(0, v)| \\ &\qquad \le b(|v| + 1) I(\nexists \text{ an optimal path } \pi \text{ from } 0 \text{ to } v \text{ with } |\pi| < c_1 |v|). \end{aligned}$$

By this inequality and condition (iv), we get immediately

$$\operatorname{Var}(\hat{T}(v)) - \operatorname{Var}(T(v)) = o\bigl(|v| / \log(|v|)\bigr),$$

so that it is sufficient to prove (12) for $\hat{T}(0, v)$.

Now, with $f = \hat{T}(0, v)$ and $\tilde{f} = \hat{T}(z, v + z)$ [with $z = z(Y)$ as in Section 4] the proof follows that of Theorem 1.6, with the following modifications:

A few lines above (28) we used that γ has at most $C_7 m$ vertices in the box $B_m(w)$. In the current situation we have to add, as a correction term, the probability that γ has more than $C_7 m$ vertices in that box. It follows easily from condition (iv) that, with a proper choice of C_7, this probability goes to 0 faster than any power of m. Hence (recalling the definition of m), it is clear that (28) remains true. Therefore, the denominator of the second term in the proof of Theorem 1.6 is, in the current situation, again larger than a constant times $\log|v|$.

A few lines before (44) we applied Observation 4.1(a) (which used the fact that all t-values were larger than some positive a) to conclude that the number of vertices of γ in a certain box of length of order k is at most some constant times k. In the current situation we do not have this strong bound, but we can obviously conclude that this number is at most the total number of vertices in the box. Because of this, the k in (44) is, in our current situation, replaced by k^d.

A few lines above (45), we again used Observation 4.1(a). Again we have to replace a factor k by k^d. By this (and the previous remark) the k^2 in (45), and therefore also in (46) becomes k^{2d}.

By the definition of $\hat{T}$, the statement about the size of the lattice animal [a few lines above (47)] still holds (with appropriate constants). By this and the earlier remarks, we now get (47) with the factor k^2 replaced by k^{2d}. By condition (i′), the denominator in the right-hand side of (48) is now of order $\exp(\varepsilon_0 k^{\varepsilon_1})$, so that the sum over k in this modified form of (49) is still finite.

This completes the proof of Theorem 1.7.

Acknowledgments. Our interest in this subject was triggered by a lecture series on first-passage percolation by Vladas Sidoravicius in the fall of 2009.

We thank Chuck Newman for referring us to certain results on greedy lattice animals.

REFERENCES

[1] BENAÏM, M. and ROSSIGNOL, R. (2008). Exponential concentration for first passage percolation through modified Poincaré inequalities. *Ann. Inst. H. Poincaré Probab. Statist.* **44** 544–573. MR2451057

[2] BENJAMINI, I., KALAI, G. and SCHRAMM, O. (2003). First passage percolation has sublinear distance variance. *Ann. Probab.* **31** 1970–1978. MR2016607

[3] COX, J. T., GANDOLFI, A., GRIFFIN, P. S. and KESTEN, H. (1993). Greedy lattice animals. I. Upper bounds. *Ann. Appl. Probab.* **3** 1151–1169. MR1241039

[4] GANDOLFI, A. and KESTEN, H. (1994). Greedy lattice animals. II. Linear growth. *Ann. Appl. Probab.* **4** 76–107. MR1258174

[5] GRAHAM, B. T. (2009). Sublinear variance for directed last-passage percolation. *J. Theor. Probab.* To appear. Available at arXiv:0909.1352.

[6] HÄGGSTRÖM, O. and NELANDER, K. (1999). On exact simulation of Markov random fields using coupling from the past. *Scand. J. Statist.* **26** 395–411. MR1712047

[7] HÄGGSTRÖM, O. and STEIF, J. E. (2000). Propp–Wilson algorithms and finitary codings for high noise Markov random fields. *Combin. Probab. Comput.* **9** 425–439. MR1810150

[8] HIGUCHI, Y. (1993). A sharp transition for the two-dimensional Ising percolation. *Probab. Theory Related Fields* **97** 489–514. MR1246977

[9] HIGUCHI, Y. and ZHANG, Y. (2000). On the speed of convergence for two-dimensional first passage Ising percolation. *Ann. Probab.* **28** 353–378. MR1756008

[10] HOWARD, C. D. and NEWMAN, C. M. (1999). From greedy lattice animals to Euclidean first-passage percolation. In *Perplexing Problems in Probability*. *Progress in Probability* **44** 107–119. Birkhäuser, Boston, MA. MR1703127

[11] JOHANSSON, K. (2000). Transversal fluctuations for increasing subsequences on the plane. *Probab. Theory Related Fields* **116** 445–456. MR1757595

[12] KAHN, J., KALAI, G. and LINIAL, N. (1988). The influence of variables on Boolean functions. In *Proc. 29th Annual Symposium on Foundations of Computer Science* 68–80. IEEE Computer Society Press, Washington, DC.

[13] KESTEN, H. (1993). On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** 296–338. MR1221154

[14] KISS, D. (2010). A generalization of Talagrand's variance bound in terms of influences. Available at arXiv:1007.0677.

[15] MARTIN, J. B. (2002). Linear growth for greedy lattice animals. *Stochastic Process. Appl.* **98** 43–66. MR1884923

[16] MARTINELLI, F. and OLIVIERI, E. (1994). Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Comm. Math. Phys.* **161** 447–486. MR1269387

[17] PROPP, J. G. and WILSON, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Struct. Algorithms* **9** 223–252. MR1611693

[18] ROSSIGNOL, R. (2008). Threshold phenomena on product spaces: BKKKL revisited (once more). *Electron. Comm. Probab.* **13** 35–44. MR2372835

[19] TALAGRAND, M. (1994). On Russo's approximate zero–one law. *Ann. Probab.* **22** 1576–1587. MR1303654

[20] VAN DEN BERG, J. (2008). Approximate zero–one laws and sharpness of the percolation transition in a class of models including two-dimensional Ising percolation. *Ann. Probab.* **36** 1880–1903. MR2440926

[21] VAN DEN BERG, J. and STEIF, J. E. (1999). On the existence and nonexistence of finitary codings for a class of random fields. *Ann. Probab.* **27** 1501–1522. MR1733157

CWI
SCIENCE PARK 123
1098 XG AMSTERDAM
THE NETHERLANDS
AND
DEPARTMENT OF MATHEMATICS
VU UNIVERSITY—FACULTY OF SCIENCES
DE BOELELAAN 1081A
1081 HV AMSTERDAM
THE NETHERLANDS
E-MAIL: J.van.den.Berg@cwi.nl

CWI
SCIENCE PARK 123
1098 XG AMSTERDAM
THE NETHERLANDS
E-MAIL: D.Kiss@cwi.nl