Large scale degrees and the number of spanning clusters for the uniform spanning tree

Itai Benjamini

October 8, 1998

Abstract

We study large scale properties of the uniform spanning tree in $\mathbb{Z}^d$. It is shown that inside the n-cube of $\mathbb{Z}^d$, there are $o(n^d)$ vertices with large scale degree 2. In 2D the number of vertices with large scale degree 3 is uniformly bounded and there are no vertices with large scale degree bigger than 3. Also, it is shown that the number of spanning clusters for the uniform spanning tree in a square is tight as the square grows, and that this is not the case in dimensions 4 and higher. A similar transition is a known conjecture for critical Bernoulli percolation.

1 Introduction

The uniform spanning tree model is getting some attention recently. In this note we would like to remark on large scale properties of the uniform spanning tree in $\mathbb{Z}^d$: large scale degrees and the number of spanning clusters in a large cube. Consider the uniform spanning tree of the square grid $\mathbb{Z}^d$, or the n-cube, denoted C_n^d, in $\mathbb{Z}^d$. Burton and Pemantle (1993) showed how to calculate the probability for a fixed edge to be in tree, as well as any other cylindrical event. In particular the distribution of the degree of a fixed vertex can be calculated. Recently Schramm (1998), constructed an object "living" in $\mathbb{R}^2$, which is conjecturally the scaling limit of the uniform spanning tree in

$\mathbb{Z}^2$. Having the proposed scaling limit in mind, here we are interested in the large scale degree of vertices. This is not a cylindrical event, and presents properties of the tree which are supposed to be inherited by the limiting object. Thus, different reasoning is used in the study of large scale degrees.

Roughly, the large scale degree of a vertex v, in a spanning tree, is the number of subtrees rooted at $v \in C_n^d$, which reach distance of order n from v in the $\mathbb{Z}^d$ graph metric.

We will show that the number of vertices with large scale degree 2 inside the n-cube is $o(n^d)$, in 2D we further show that the number of vertices with large scale degree 3, is uniformly bounded and that with probability going to 1, as n grows, there are no vertices with large scale degree bigger than 3. See below for the precise result.

We now turn to spanning clusters. Consider again C_n^d, the n-cube in $\mathbb{Z}^d$, i.e. the vertex set is $\{(x_1, ..., x_d) \in \mathbb{Z}^d \mid 0 \leq x_i \leq n\}$, and edges are between any two vertices with Euclidean distance 1. Assign each edge in the cube the value 0 or 1. A spanning cluster is a connected component in the $0, 1$ configuration consisting only of 1's, which connect the face $x_1 = 0$ to its parallel face, $x_1 = n$. In particular, in our discussion, an edge is assigned 1 iff it is in the spanning tree. We would like to remark on the number of spanning clusters for the uniform spanning tree in $\mathbb{Z}^d$. Our motivation to look at this problem comes from percolation. We will see that the analogue of the conjectural picture for spanning clusters of critical Bernoulli percolation is in a larger part provable for the uniform spanning tree. It is conjectured that for critical Bernoulli percolation in a large cube, for $d \leq 6$, the number of disjoint spanning clusters is tight, as the size of the cube grows. Yet for $d > 6$, the expected number of spanning clusters grows to infinity, with the size of the cube. By the BK inequality the 2D case follows; the rest is still open. (See for instance Borgs, Chayes, Kesten and Spencer (1998), Aizenman (1997) and Cardy (1997)). The point of this note is that the qualitative transition between low and high dimensions is established for the uniform spanning tree (the exact picture is still open, in particular tightness in 3D). This can serve as additional support for the Bernoulli percolation conjecture.

2 Some background and formulation of the results

The uniform spanning tree measure for a finite connected graphs is just the uniform measure on the space of spanning trees of the graph. The uniform spanning tree measures for infinite graphs are weak limits of uniform spanning trees from finite subgraphs. See Pemantle (1991) or Benjamini, Lyons, Peres and Schramm (1998) for an introduction to and a study of uniform spanning tree measures.

Häggström (1995) showed that the uniform spanning tree measures arise as limits of (Fortuin-Kasteleyn) random cluster measures. Such measures generalize ordinary Bernoulli percolation, Ising and Potts models of statistical physics. Many questions that are difficult and often still unsolved for some of these models, become tractable for the uniform spanning tree model because of its close connection to potential theory and random walks. Conformal invariance as well as other properties, such as the distribution of spanning clusters discussed here, suggests that the uniform spanning tree behave like a critical model, although there is no parameter in the model.

Wilson's algorithm. Let G be a finite connected graph. Every connected graph has a spanning tree. It is not obvious how to choose one uniformly at random, and several sophisticated algorithms have been devised to do this. The fastest algorithm known, which we describe below, and then use, is due to Wilson (1996). To describe Wilson's algorithm, we define the loop erasure of a path. If γ is any finite path $\{v_0, v_1, ..., v_l\}$ in G, we define the loop erasure of γ, denoted $LE(\gamma)$, by erasing cycles in γ in the order they appear. A slightly different, and more precise, inductive description of the loop erasure $LE(\gamma) = \{u_0, u_1, ..., u_m\}$ of a path $\gamma = \{v_0, v_1, ..., v_l\}$ is as follows. The first vertex u_0 of $LE(\gamma)$ is the first vertex v_0 of γ. Suppose that u_j has been set. Let k be the last index such that $v_k = u_j$. Set $u_{j+1} := v_{k+1}$ if $k < l$; otherwise, let $LE(\gamma) := (u_1, ..., u_j)$. Note that $LE(\gamma)$ is still well defined when γ is an infinite path that visits no vertex infinitely often.

In order to generate a random spanning tree, first pick any vertex r to be the "root" of the tree. Then create a growing sequence of trees t_i $(i \geq 0)$ as follows. Choose any ordering $\{v_1, ..., v_n\}$ of the vertices. Let $t_0 := \{r\}$. Suppose that the tree t_i has been generated. Start an independent simple random walk at v_{i+1} and stop at the first time it hits t_i. (If $v_{i+1} \in t_i$, then the random walk will consist only of $\{v_{i+1}\}$.) Now create t_{i+1} by adding to t_i the loop erasure of this random walk. Then t_{i+1} is a tree. The output of Wilson's algorithm is the set of edges of the tree $T = t_n$. (The root is forgotten.)

Wilson (1996) proved that if G is a finite graph, then Wilson's method yields a uniform spanning tree.

Boundary conditions. In the study of spanning clusters for the cube, one can consider several uniform spanning tree measures. The most interesting is as follows. Pick a tree according to the uniform spanning tree measure in all of $\mathbb{Z}^d$, and look at the restriction to the n-cube. This measure will be used in the analysis of the number of spanning clusters in dimensions $d \geq 4$. Another natural measure is start with a finite n-cube, ignoring the rest of the lattice, and pick a spanning tree uniformly, this is the free uniform spanning tree. For the study of spanning clusters in 2D, we will use a variant of this measure, that is, we will consider the uniform measure on spanning trees of the n-square, when all the vertices on the right side are identified (wired) to single vertex. Call this measure, right wired. In this setup, two spanning clusters are disjoint if they intersects only in the rightmost wired vertex. The study of spanning clusters for the $\mathbb{Z}^2$ uniform spanning tree measure restricted to the square (the first measure above), is much harder. Regarding one possible approach in this setup, see the remark after the proof of theorem 1.

With these measures in mind, we have,

Theorem 1 *(number of spanning clusters) For $d = 2$, with the right wired measure, the number of spanning clusters is tight as the size of the square grows. For $d \geq 4$, the expected number of spanning clusters grows to infinity with the size of the cube.*

Remarks:

1) We can't decide the case $d = 3$, but **conjecture** tightness to hold.

2) Pemantle (1991), based on work by Lawler on loop erased random walk, proved that the uniform spanning forest of $\mathbb{Z}^d$ is a tree a.s., iff $d \leq 4$. Thus, the case $d = 4$, is interesting, since there is a unique infinite cluster, yet many spanning clusters.

3) For critical percolation in 2D, Aizenman (1997) proved that the probability of having k vertex disjoint crossings is bounded between $Ae^{-\alpha k^2}$ and $e^{-\alpha' k^2}$, for some constants A, α and α'. Cardy (1997) predicts that the exact exponent is $-(2\pi/3)k(k-1/2)$. Duplantier (private communication) predicts the exponent for the uniform spanning tree to be $-(\pi/2)k(k-1)$.

4) Consider the uniform spanning tree in $\mathbb{Z}^2$. Fix $c > 0$. What is the limit as n grows of the probability, that there exists a spanning cluster for the $n \times cn$-rectangle? Cardy (1992) proposed an answer to this for critical Bernoulli percolation.

5) Recently Lawler (1998) gave a lower bound on the growth exponent for loop-erased random walk in two dimension, which is strictly bigger than 1. Aizenman and Burchard (1998), studied regularity properties of random curves, obtained as scaling limits. The theorem above and simple crossings bounds obtained by Wilson's algorithm provide the analogue of the "Russo-Seymour-Welsh theory" needed in order for their work to apply.

For a study of the large scale degrees, we will consider the $3n$-square in Z^2, with free boundary conditions. And we will define large scale degree, as follows:

Definition. Consider the n-square centered in the $3n$-square. The large scale degree of a vertex v, in the n-square, denoted $lsd(v)$, is the maximal number of disjoint paths in the tree that connect v to the boundary of the $3n$-square.

In this setup, we have

Theorem 2 *(Large scale degree)*

1. $\#\{v \in C_n^d \mid lsd(v) > 1\} = o(n^d)$ *in probability, and*
$$E\Big(\#\{v \in C_n^2 \mid lsd(v) > 1\}\Big) < Cn^{4/3}\log n.$$
2. $E(\#\{v \in C_n^2 \mid lsd(v) > 2\})$ *is uniformly bounded in* n.
3. $P(\exists v \in C_n^2, lsd(v) > 3)$ *goes to* 0*, with* n.

Remark:

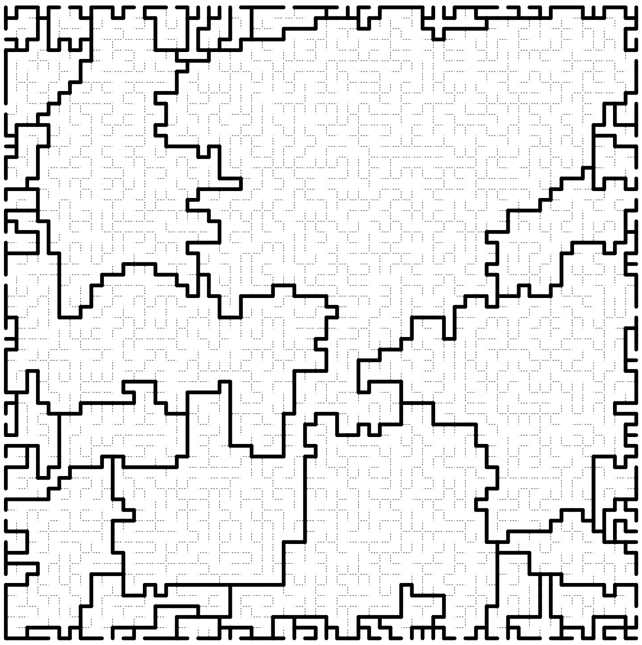

Figure 1: One picture (by Oded) is better than a thousand words (by ...).

1) See figure 1 for an illustration of the theorem.

2) The 4/3 in item 1. should be improved to 5/4, which is believed to be the growth exponent for the loop-erased random walk (see Duplantier (1992)). Item 2. might be true also in dimension 3, and maybe item 3. holds in any dimension.

3) Since the number of vertices v with $lsd(v) > 2$ is uniformly bounded in n, it's limit distribution, if it exists, might serve as a natural conformal invariant for the annulus.

3 Proofs and further comments

We start with spanning clusters.

Proof of Theorem 1. We start with $d = 2$. Suppose that there is at least one crossing, and condition on the upper left-right crossing. The configuration of the uniform spanning tree in the domain below this crossing is independent from the configuration above, and is just the uniform spanning tree for the domain wired at the upper crossing and the right boundary. With probability at least $1/2$, simple random walk starting at the bottom left corner of the square, will hit the upper crossing, before hitting the right side of the square. Therefore, by Wilson's algorithm, with probability at least $1/2$ there is a path in the spanning tree, connecting the bottom left vertex to the upper crossing, which is disjoint from the right wired side. On this event, there is just one left-right crossing. By each time conditioning on the next upper crossing, for the remaining domain, and by the argument above, we get that the probability of having k disjoint crossings is bounded by 2^{-k}.

Remark: Note that if you consider k independent simple random walks starting equally spaced on the left side of the square, then there is $c > 0$, so that with probability bounded below by c^{-k^2}, the k simple random walks will reach the right side without any intersection. By Wilson's algorithm, this gives a lower bound for the probability of having k disjoint spanning clusters.

We turn now to $d \geq 4$. As was pointed out in Benjamini, Lyons, Peres and Schramm (1998) Thm 5.1, the uniform spanning tree measure for $Z^d, d \geq 3$, can be constructed via Wilson's method rooted at infinity, that is, start simple random walk at some vertex, and loop erase the infinite path. Then pick a vertex not on the path. Start simple random walk from the new vertex till it hits the first path or the root (infinity) and loop erase, and so on. So the case $d \geq 4$, for that measure, follows via Wilson's algorithm from standard properties of simple random walk in $Z^d, d \geq 4$. See Lawler (1991), for the fact that the probability that 2 and hence any k paths of k independent simple random walks, starting distance n apart will intersect before reaching distance Cn (from one of the starting vertices), is going to 0, for any fixed $C > 0$, as n grows. In dimension 4, the probability that two random walks starting at distance n apart intersect before reaching distance Cn is of order $c/\log n$. In Benjamini, Kesten, Peres and Schramm (1998), it is shown that in $d > 4$, the expected size of the tree containing 0, in the uniform spanning forest, inside a ball of radius n, is of order Cn^4. Thus inside the n-cube one expects about n^{d-4}, large connected components, and hence, up to a constant, as many spanning clusters. $\square$

Note that if we work with a measure as in the 2D, wired on the "right" face and ignoring the rest of the lattice, then same proof works.

How many spanning clusters are there in 4D?

Remark: Benjamini, Lyons, Peres and Schramm (1998) conjectured the following BK-Type Inequality: Denote by E the set of edges in a finite graph. We say that $A, B \subset 2^E$ occur disjointly for $F \subset E$ if there are disjoint sets $F_1, F_2 \subset E$ such that $F' \in A$ for every F' with $F' \cap F_1 = F \cap F_1$ and $F' \in B$ for every F' with $F' \cap F_2 = F \cap F_2$. Let $A, B \subset 2^E$ be increasing. Then the probability that A and B occur disjointly for the random spanning tree T is at most $P[T \in A]P[T \in B]$. (The BK inequality of van den Berg and Kesten says that the same is true when T is a random subset of E chosen according to any product measure on 2^E, see Grimmett (1989).)

If true, this will allow to remove the wired boundary, and to consider the uniform spanning tree of $\mathbb{Z}^2$, restricted to a the n-square. Also it will help in getting better tail estimates.

proof of Theorem 2.

The first part of 1. follows from the fact that any tree in the uniform spanning forest of $\mathbb{Z}^d$, has a.s. one end. (see Pemantle for dimensions $2 \leq d \leq 4$, and Benjamini, Lyons, Peres and Schramm (1998), for a simple proof in 2D and for all other dimensions). In particular for a fixed v the probability that v is in the middle of a path of length m, in the tree, goes to 0, with m. The uniform spanning tree measure on $\mathbb{Z}^d$, is a weak limit of the free uniform spanning tree measure on the n-cube. So first pick m so that in the uniform spanning tree of $\mathbb{Z}^d$, the origin has chance less than ϵ of connecting the origin in 2 disjoint paths to distance $m/2$, and then freeze m and pick n so that the chance of the cylinder event above in the free spanning tree in the n-cube is smaller than 2ϵ.

The sharper 2D case follows from duality. The dual of free spanning tree on the $3n$-square is the wired spanning tree. See Benjamini, Lyons, Peres and Schramm (1998), for a discussion of the duality and applications. If $lsd(v) > 1$ then in the dual tree, on the dual lattice, there should be two vertices at distance 1 (neighbours of v in the dual) that are connected in the wired spanning tree only through the boundary; But by Lawler's (1991) upper bound on intersection probabilities, via Beurling's projection theorem, which can be used here, this has probability bounded by $n^{-2/3} \log n$; multiplying by n^2 we get item 1.

Note that, since the set of vertices with $lsd(v) = 2$ contains a loop-erased path, which get to the boundary of the $3n$-square twice, we get that the number of vertices on such a path, inside the n-square, is bounded by $n^{4/3} \log n$, recovering another result in Lawler (1991).

We move to item 2. In theorem 1, the exponential bound on the number of crossings was proved for a square with the right side wired, here we need a variant regarding crossings of an annulus, between the n-square and the $3n$-square, wired at the boundary of the $3n$-square. Following the argument in the proof of the first part of theorem 1 we get that the probability for k crossings from, say, the top side of the n-square to the $3n$-square decays exponentially. Yet P(there are k crossings of the annulus) $\leq$ $4P$(there are $k/4$ crossings from the top side).

Now if there are k vertices in the n-square with $lsd(v) > 2$, then by the tree structure there are at least $k+1$ disjoint crossings from the boundary of the n-square, to the boundary of the $3n$-square. Again by the duality between the free and the wired, there are $k+1$ or more disjoint crossings in the dual wired spanning tree, connecting the boundary of the n-square to the wired boundary of the $3n$-square. By annulus version of the first part of theorem 1, the probability of having such k crossings decays exponentially with k.

Remark: Note that as n grows the mean number of vertices with $lsd(v) > 2$ is bounded away from 0. Using Wilson's algorithm it is not hard to see that with probability bounded away from 0, with n, the meeting vertex for the tree spanned by the mid-vertex on the top side of the $3n$-square, and the two mid-vertices on the left and right sides, will be inside the middle n-square.

To prove 3, start growing the tree in the $3n$-square via Wilson's algorithm. Assume at some stage a vertex v, with $lsd(v) = 3$, was created. Part of the boundary of the $3n$-square can no longer connect to v, to increase it's large scale degree. For any vertex on the $3n$-boundary for which simple random walk can in principle hit v, before hitting the rest of the tree, the probability this will happen is going to zero, with n, by the discrete version of Beurling's projection theorem (see Kesten (1987)). Pick u_1 on the boundary, to be the middle vertex of the part of the boundary that can hit the tree at v. Assume simple random walk, starting at u_1, failed to hit the tree at v. Pick u_2 as the middle of the boundary part, that can still hit the larger tree at v, and define u_i inductively in the same way, by each time picking the middle vertex of the part of the boundary that still can hit, the previously constructed tree at v. Let p_i, be the probability simple random walk reach distance n from u_i, without hitting the part of the boundary, that can no longer "connect" to v. The p_i's decay exponentially with i, with a rate that is independent of n. By Beurling's projection the p_i's bounds from above the probability that simple random walk starting at u_i, will hit the tree, constructed so far, at v. We get that the probability that the large scale degree of v, will increase to 4 is going to 0, with n. Since the probability that the first trails will hit v, is going to zero with n, and the probability the later trails will succeed is small by the uniform exponential decay. By 2. the number of vertices with $lsd(v) = 3$ is tight, and 3. follows. $\square$

Acknowledgements: I would like to thank Bertrand Duplantier, Balint Toth and Wendelin Werner for useful discussions, Yuval Peres and Oded Schramm for simplifying some of the proofs and David Wilson for remarks on a preliminary version.

References

[1] M. Aizenman, On the number of incipient spanning clusters, *Nucl. Phys.* B[FS] **485**, 551, (1997)

[2] M. Aizenman and A. Burchard, Hölder Regularity and dimension bounds for random curves, preprint, (1998)

[3] I. Benjamini, H. Kesten, Y. Peres and O. Schramm, Paper in preparation, (1998)

[4] I. Benjamini, R. Lyons, Y. Peres and O. Schramm, Uniform spanning forests, preprint, (1998)

[5] C. Borgs, J. Chayes, H. Kesten and J. Spencer, Uniform boundedness of critical crossing probabilities implies hyperscaling, preprint, (1998)

[6] R. Burton and R. Pemantle, Local characteristics and limit theorems for spanning trees and domino tilings via transfer-impedances, *Ann. Prob.* **21**, 1329–1371, (1993)

[7] J. Cardy, Critical percolation in finite geometries, *J. Phys. A* **25**, 201–206, (1992)

[8] J. Cardy, The number of incipient spanning clusters in two-dimensional percolation, cond-mat/9705137, (1997)

[9] B. Duplantier, Loop-erased self-avoiding walks in 2D, *Physica A* **191**, 516–522, (1992)

[10] G. Grimmett, *Percolation*, Springer-Verlag, (1989)

[11] O. Häggström, Random-cluster measures and uniform spanning trees, *Stoch. Proc. Appl.* **59**, 267–275, (1995)

[12] H. Kesten, Hitting probabilities of random walks on $\mathbb{Z}^d$, *Stoc. Proc. and Appl.* **25**, 165–184, (1987)

[13] G. Lawler, *Intersections of Random Walks.* Brikhäuser, (1991)

[14] G. Lawler, A lower bound on the growth exponent for loop-erase random walk in two dimension, preprint, (1998)

[15] R. Pemantle, Choosing a spanning tree for the integer lattice uniformly, *Ann. Prob.* **19**, 1559–1574, (1991)

[16] D. Wilson, Generating random spanning trees more quickly than the cover time, *1996 ACM sympos. Theory of Computing*, 296–303, (1996)

[17] O. Schramm, Scaling limits of loop-erased random walks and random spanning trees, preprint, (1998)

Itai Benjamini,
Weizmann Institute,
Rehovot,
Israel 76100

itai@wisdom.weizmann.ac.il
http://www.wisdom.weizmann.ac.il/~itai/