A homological upper bound on critical probabilities for hyperbolic percolation

Nicolas Delfosse* and Gilles Zémor†

August 19, 2014

Abstract

We study bond percolation for a family of infinite hyperbolic graphs. We relate percolation to the appearance of homology in finite versions of these graphs. As a consequence, we derive an upper bound on the critical probabilities of the infinite graphs.

1 Introduction

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an infinite connected graph. Every edge is declared to be *open* with probability p, otherwise it is *closed*. This endows subsets of edges with a product probability measure by declaring edges to be open or closed independently of the others, and creates a random open subgraph ε. If a given edge e belongs to an infinite connected component of ε, we say that (bond) percolation occurs. If the graph $\mathcal{G}$ is edge-transitive, then the probability of percolation does not depend on e and we may denote this probability by $f(p)$. Arguably, the most studied parameter of percolation theory is the *critical probability* $p_c = p_c(\mathcal{G})$ which is the supremum of the set of p's for which $f(p) = 0$.

Ever since the seminal work of Kesten [13] percolation was extensively studied on the lattices associated to $\mathbb{Z}^d$, for background see [9]: in the present paper, we are interested in percolation on regular tilings of the hyperbolic plane. This topic was first introduced by Benjamini and Schramm [3], and further studied in [4, 10, 1] among other papers. Specifically, our focus is on the family of graphs that we shall denote by $G(m)$, for $m \geq 4$, that are regular of degree m, planar, and tile the plane by elementary faces of length m. For $m = 4$, the graph $G(m)$ is exactly the square $\mathbb{Z}^2$ lattice. The local structure of the graph $G(5)$ is represented in Figure 1.

Our goal is to study the critical probabilities of these lattices. The simple lower bound $1/(m-1) \leq p_c$ can be derived since $1/(m-1)$ is the critical probability for the m-regular tree, and our main concern here is on dealing with upper bounds. Critical probabilities for hyperbolic tilings were studied numerically by Baek et al. [1] and also by Gu and Ziff [10] who obtain a "Monte Carlo" upper bound $p_c < 0.34$ for $G(5)$. In previous work by the present authors [7], the rigorous upper bound $p_c < 0.38$ was obtained for $G(5)$ as a by-product of the study of the erasure-correcting capabilities of a family of quantum error-correcting

*Département de Physique, Université de Sherbrooke, Sherbrooke, Québec, J1K 2R1 Canada, nicolas.delfosse@usherbrooke.ca

†Institut de Mathématiques de Bordeaux, UMR 5251, université de Bordeaux, zemor@math.u-bordeaux.fr

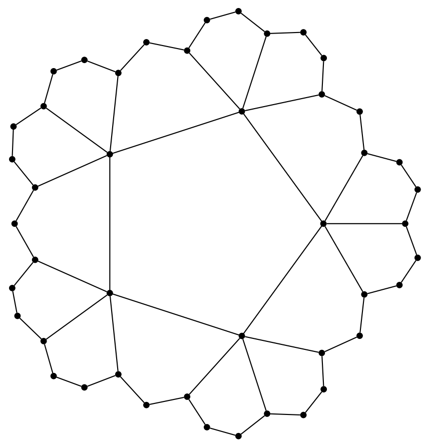

Figure 1: The local structure of the graph $G(5)$

codes. In the present paper we shall obtain a substantially improved upper bound on critical probabilities that gives $p_c < 0.30$ for $G(5)$.

We remark that we restrict ourselves to the hyperbolic tilings $G(m)$ because they are self-dual and our method is better suited for this case, but results on the critical probabilities for the self-dual case can lead to results for the general case [14].

Classically, one uses finite portions of the infinite graph $\mathcal{G}$ to devise intermediate tools for studying percolation. For example, in the original $\mathbb{Z}^2$ setting, the standard (by now) method that leads to the computation $p_c = 1/2$ is to consider $n \times n$ finite grids and study the probability of the appearance of an open path linking the south boundary to the north boundary (or east to west) [9]. In the hyperbolic setting however, trying to mimic this approach directly quickly leads to serious obstacles: what finite portion of the infinite graph $G(5)$ (say) should one consider, and which parts of the boundary should be matched when looking for the appearance of finite open paths ? We shall overcome this difficulty by appealing to finite graphs $G_t(m)$ that are everywhere locally isomorphic to $G(m)$, meaning that every ball of radius t of $G_t(m)$ is required to be isomorphic to a ball of radius t in the infinite graph $G(m)$. We shall derive an upper bound $p_c \leq p_h$ on the critical probability by defining a quantity p_h such that, when $p > p_h$, then with probability tending to 1 when t tends to infinity, $G_t(m)$ must contain an open cycle that can not be expressed as a sum modulo 2 of elementary faces. Our end result will be an expression for the upper bound p_h that involves only the structure of the infinite graph $G(m)$, but the existence of the finite graphs $G_t(m)$ (which is non-obvious) will be crucial to the derivation of p_h.

Outline and results: Sections 2 and 3 are background. In Section 2 we give a short description of a construction of the graphs $G_t(m)$ due to Širáň. We shall need to consider the cycles of those graphs that are not expressible as sums of faces, i.e. that are homologically non-trivial: we shall therefore need background on homology that is dealt with in Section 3.

In Section 4 we study the appearance of homology in random subgraphs of the finite graphs $G_t(m)$. We introduce a crucial quantity $D(p)$ that we name the *rank difference function* and that captures the limiting behaviour of the difference of the dimensions of the homologies of the two random subgraphs of $G_t(m)$ chosen through the parameters p and $1 - p$. We then define the quantity

$$p_h = \sup\left\{p, p - \frac{2}{m} + D(p) = 0\right\}.$$

The main result of this section, Theorem 4.5, is that p_h is an upper bound on the critical probability of $G(m)$. We actually conjecture that for $m \geq 5$ (i.e. the genuinely hyperbolic, or non-amenable, case) this upper bound is also a lower bound, i.e. $p_c = p_h$. This would show that for these graphs the critical probability is local in a sense close to [2]. That $p_c \leq p_h$ was derived in [7] in a roundabout way, through the study of the erasure-decoding capabilities of quantum codes associated to the tilings $G_t(m)$. The present proof not only removes the reference to quantum coding, it is intrinsically shorter and more direct.

Section 5 is dedicated to finding an explicit expression for the rank difference function $D(p)$, and hence for the upper bound p_h. Our main result is Theorem 5.3, which expresses $D(p)$ as the series:

$$D(p) = \frac{2}{m}\sum_C \left(\frac{1}{|V(C)|}\left(p^{|E(C)|}(1-p)^{|\partial(C)|} - (1-p)^{|E(C)|}p^{|\partial(C)|}\right)\right), \tag{1}$$

where C ranges over all connected subgraphs of $G(m)$ containing a given vertex, where $V(C), E(C)$ denote the vertex and edge set of C, and where $\partial(C)$ denotes the set of edges with at least one endpoint in C, which are not in $E(C)$. As mentioned, this expression for $D(p)$ does not involve the graphs $G_t(m)$ anymore, but its proof crucially relies on their existence.

Section 6 proves that replacing $D(p)$ in (1) by a truncated series continues to yield an upper bound on the critical probability p_c of $G(m)$ (Theorem 6.1). This allows us to compute explicit numerical upper bounds on p_c. Finally, Section 7 summarizes the results with Theorem 7.1 and gives some concluding comments.

2 Finite quotient of the regular hyperbolic tilings

We are unaware of any method for constructing the required finite versions of $G(m)$ that does not involve a fair amount of algebra. In this section, we briefly recall Širáň's method to construct such finite versions of the regular hyperbolic tiling $G(m)$. The first step is to construct $G(m)$ from a group of matrices over a ring of algebraic integers. Then this group is reduced modulo a prime number to yield the desired finite graph.

Denote by $P_k(X) = 2\cos(k \arccos(X/2))$ the k-th normalized Chebychev polynomial and let $\xi = 2\cos(\pi/m^2)$. Let $m \geq 5$ and consider the group $T(m)$ generated by the two following matrices of $SL_3(\mathbb{Z}[\xi])$.

$$a = \begin{pmatrix} P_m(\xi)^2 - 1 & 0 & P_m(\xi) \\ P_m(\xi) & 1 & 0 \\ -P_m(\xi) & 0 & -1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} -1 & -P_m(\xi) & 0 \\ P_m(\xi) & P_m(\xi)^2 - 1 & 0 \\ P_m(\xi) & P_m(\xi)^2 & 1 \end{pmatrix}.$$

The group $T(m)$ admits the presentation

$$T(m) = \langle a, b \mid a^m = b^m = (ab)^2 = 1\rangle. \tag{2}$$

With this group we associate its *coset graph*. The coset graph associated with (2) is defined to be the infinite planar tiling whose vertex set, respectively edge set and face set, is the set of left cosets of the subgroup $\langle a \rangle$, respectively the set of left cosets of the subgroup $\langle ab \rangle$ and the subgroup $\langle b \rangle$. A vertex and an edge, or an edge and a face, are incident if and only if the corresponding cosets have a non-empty intersection.

For example, the coset $\langle a \rangle = \{1, a, a^2, \dots, a^{m-1}\}$ defines a vertex of the graph $G(m)$ and is incident to the m edges represented by the cosets

$$\langle ab \rangle,\ a\langle ab \rangle,\ a^2\langle ab \rangle,\ \dots,\ a^{m-1}\langle ab \rangle.$$

We can see that the coset graph is m-regular and that its faces contain m edges. It is straightforward to check that the coset graph associated with (2) is the infinite planar graph $G(m)$ [15].

The basic idea to derive a finite version of this tiling is to reduce the matrices defining the group $T(m)$ modulo a prime number. We can reduce the coefficients of the matrices of $T(m)$ thanks to the ring isomorphism $\mathbb{Z}[\xi] \simeq \mathbb{Z}[X]/h(X)$, where $h(X) \in \mathbb{Z}[X]$ is the minimal polynomial of the algebraic number ξ. This induces a ring morphism $\pi_p : SL_3(\mathbb{Z}[\xi]) \to SL_3(\mathbb{F}_p[X]/\bar{h}(X))$ where $\bar{h}(X)$ is the reduction modulo p of the polynomial $h(X)$. Denote by $\bar{T}^p(m)$ the image of the group $T(m)$ by the morphism π_p. The coset graph associated with the group $\bar{T}^p(m)$ is defined from the cosets of $\bar{T}^p(m)$, exactly like the coset graph of $T(m)$.

Širáň proved that for a well chosen family of prime numbers p, this construction provides a sequence of finite tilings $(G_t(m))_t$ which is locally isomorphic to the infinite tiling $G(m)$ [15]. Precisely:

Theorem 2.1. *For every integer $m \geq 5$, there exists a family of finite tilings $(G_t(m))_{t \geq m}$ and some constant K such that every ball of radius t of $G_t(m)$ is isomorphic to every ball of radius t in $G(m)$. Furthermore, the number of vertices of $G_t(m)$ is at most K^t.*

By construction, the graphs $G_t(m)$ are vertex transitive. Indeed, each element of the group $\bar{T}(m)$ induces a graph automorphism of the coset graph by left multiplication. An automorphism which sends a vertex $x\langle a \rangle$ onto the vertex $y\langle a \rangle$ is given by the left multiplication by yx^{-1} of the cosets representing the vertices. For the same reason, $G_t(m)$ is also edge-transitive and face-transitive.

To be sure that the faces of the graph $G_t(m)$ are not degenerate, we require $t \geq m$. We will also use the fact that G_t is a self-dual graph. This is a consequence of the local structure of the graph: every vertex has degree m and every face has length m.

3 Background on homology

3.1 Homology of a tiling of surface

A *tiling of a surface* is a graph cellularly embedded in a smooth surface. For us only the combinatorial structure of the surface plays a role, therefore a face of the tiling is represented as the set of edges on its boundary. We denote by $G = (V, E, F)$ such a tiling, where F is the set of faces that, as far as homology is concerned, can be thought of simply as a privileged set of cycles of the graph (V, E). With a tiling of a surface, we associate a *dual tiling* $G^* = (V^*, E^*, F^*)$. The vertices of this dual tiling are given by the faces of G. Two vertices of G^* are joined by an edge if the corresponding faces of G share an edge. Since every edge of E belongs to exactly two faces of F, there is a one-to-one correspondence between edges of G and edges of G^*. Finally, for every vertex v of V the set of edges of E incident to v defines a face of F^* through the above correspondence between E and E^*. We assume the graph and its dual have neither multiple edges nor loops. We shall also use G to refer indifferently to the graph (V, E) and to the associated tiling (V, E, F).

In the remainder of this section, we consider only finite tilings, and we order the three sets V, E and F by $V = \{v_1, v_2, \ldots, v_{|V|}\}$, $E = \{e_1, e_2, \ldots, e_{|E|}\}$ and $F = \{f_1, f_2, \ldots, f_{|F|}\}$. The *incidence matrix* of the graph (V, E) is defined to be the matrix $B(G) = (b_{ij})_{i,j}$ of $\mathcal{M}_{|V|,|E|}(\mathbb{F}_2)$ such that $b_{ij} = 1$ if the vertex v_i is incident to the edge e_j, and $b_{ij} = 0$ otherwise.

To emphasize the $\mathbb{F}_2$-linear structure of some subsets of V, E and F, we introduce the spaces of *i-chains* C_i:

$$C_0 = \bigoplus_{v \in V} \mathbb{F}_2 v, \quad C_1 = \bigoplus_{e \in E} \mathbb{F}_2 e, \quad C_2 = \bigoplus_{f \in F} \mathbb{F}_2 f.$$

In other words, the space $C_0 = \{\sum_v \lambda_v v \mid \lambda_v \in \mathbb{F}_2\}$ is the set of formal sums of vertices. The sets C_1 and C_2 are defined similarly. These chain spaces are equipped with two $\mathbb{F}_2$-linear mappings $\partial_2 : C_2 \to C_1$ and $\partial_1 : C_1 \to C_0$ defined by $\partial_2(f) = \sum_{e \in f} e$ and $\partial_1(e) = \sum_{u \in e} u$. These mappings are called *boundary maps*.

A subset of the vertex set, respectively the edge set or the face set, can be regarded as its indicator vector in C_0, respectively C_1 or C_2. This yields one-to-one correspondences between subsets and vectors, which allow us to interpret geometrically the boundary maps. In subset language, the map ∂_2 sends a subset of faces onto the set of edges on its boundary in the standard sense, and the map ∂_1 sends a subset of edges onto its "endpoints" which should be understood modulo 2, i.e. the set of vertices incident to an odd number of edges in the subset.

The singletons $\{v_i\}$, respectively $\{e_i\}$ and $\{f_i\}$, form a basis of the space C_0, respectively C_1 and C_2. The matrix of the map ∂_1 in these singleton bases is equal to the incidence matrix $B(G)$ of the graph (V, E) and the matrix of the map ∂_2 is equal to the transpose of the incidence matrix $B(G^*)$ of (V^*, E^*).

We can easily prove that the composition of these applications is $\partial_1 \circ \partial_2 = 0$, implying the inclusion $\operatorname{Im} \partial_2 \subset \operatorname{Ker} \partial_1$. We can now introduce the $\mathbb{F}_2$-homology of tilings of surfaces.

Definition 3.1. *The* first homology group *of a finite tiling of a surface G, denoted $H_1(G)$, is the quotient space*

$$H_1(G) = \operatorname{Ker} \partial_1 / \operatorname{Im} \partial_2.$$

Note that $H_1(G)$ is also an $\mathbb{F}_2$-vector space. The vectors of $\ker \partial_1$ are called *cycles.* They correspond to the subsets of edges that meet every vertex an even number of times. The set $\ker \partial_1$ of cycles of a graph is an $\mathbb{F}_2$-linear space that we refer to as the *cycle code* of the graph. The vectors of $\operatorname{Im} \partial_2$ are called *boundaries* or sums of faces and they describe the sets of edges on the boundary of a subset of F.

In what follows, we shall study the dimension of the homology group of different tilings of surfaces. The following well known property (see e.g. [5] for a proof) is used repeatedly .

Lemma 3.2. *The dimension of the cycle code of a graph $G = (V, E)$ composed of κ connected components, is $|E| - |V| + \kappa$.*

Figure 2(a) represents a square lattice of the torus. A cycle of trivial homology is drawn on Figure 2(b). This cycle is clearly a sum of faces. Two examples of cycles with non trivial homology are given in Figure 2(c) and (d). The first homology group of this tiling of the torus is a binary space of dimension 2. It is generated, for example, by an horizontal cycle which wraps around the torus, such as the one in Figure 2(c) and a vertical cycle which wraps around the torus. The cycle of Figure 2(d) is equivalent to the sum of these horizontal and vertical cycles, up to a sum of faces.

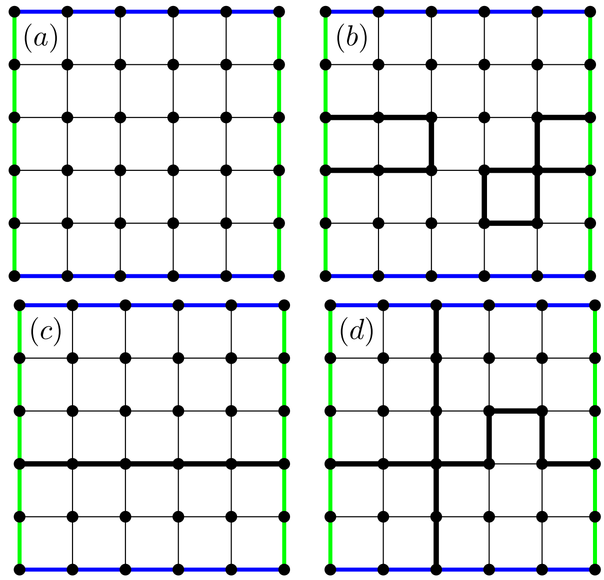

Figure 2: (a) A square tiling of the torus. The opposite boundaries are identified. (b) A cycle which is a boundary. (c) A cycle which is not a boundary. (d) A cycle which is not a boundary.

3.2 Induced homology of a subtiling

Percolation theory deals with random subgraphs of a given graph. In what follows, we introduce the homology of a subgraph of a given tiling G.

The subgraphs that we consider are obtained by selecting a subset of edges. Denote by $G = (V, E, F)$ a tiling of surface and let us consider the subgraph G_ε of G whose vertex set is exactly V and whose edge set is a given subset ε of E. This graph is not immediately endowed with a set of faces and with a homology group. The proper notion of homology for our purpose is obtained by considering the boundaries of the tiling G which are included in the subgraph G_ε. More precisely, the subset of edges ε defines the subspace $C_0^\varepsilon = C_0$, the subspace C_1^ε of C_1 made up of all formal sums of edges of ε, and the subspace C_2^ε of C_2 made up of all those vectors of C_2 whose image under ∂_2 is included in C_1^ε. The mappings ∂_1^ε and ∂_2^ε are defined as the restrictions of ∂_1 and ∂_2 to C_1^ε and C_2^ε.

Definition 3.3. *Let $G = (V, E, F)$ be a tiling of a surface and let $\varepsilon \subset E$. The* induced homology group *of G_ε is the quotient space*

$$H_1(G_\varepsilon) = \operatorname{Ker} \partial_1^\varepsilon / (\operatorname{Im} \partial_2^\varepsilon).$$

For more detailed background on the homology of surfaces and their tilings see [12, 8].

4 Appearance of homology in a random subgraph of G_t

4.1 Homology of a subgraph

This section is devoted to the analysis of the induced homology of a subgraph of $G_t(m)$. To lighten notation we omit the indices m and t and write $G = G_t(m)$. Following the notation of Section 3.2, ε denotes a subset of E and G_ε denotes the subgraph of G induced by ε.

The decomposition of the graph G_ε into connected components induces a partition of the edges of ε: the set ε is the disjoint union of the subsets $\varepsilon_i \subset E$, for $i = 1, 2, \ldots, r$ and where each set ε_i is the edge set of a connected component of G_ε. The following lemma proves that this decomposition of the graph G_ε induces a decomposition of its homology group.

Lemma 4.1. *Let $\varepsilon = \cup_{i=1}^r \varepsilon_i$ be the partition of ε derived from the decomposition of the graph G_ε into connected components. Then, the first homology group of G_ε can be decomposed as*

$$H_1(G_\varepsilon) = \bigoplus_{i=1}^r H_1(G_{\varepsilon_i}).$$

Proof. It suffices to remark that we have such a decomposition for the chain spaces $C_1^\varepsilon, C_2^\varepsilon$, the cycle code $\ker \partial_1^\varepsilon$ and the boundary space $\operatorname{Im} \partial_2^\varepsilon$ of the graph G_ε. □

The next lemma proves that if ε is composed of small clusters, then it covers no homology.

Lemma 4.2. *Let G_ε be a connected subgraph of $G = G_t(m)$. If ε contains at most t edges, then we have $H_1(G_\varepsilon) = \{0\}$.*

Proof. Since G_ε is connected and contains less than t edges, it is included in a ball of radius t. From Theorem 2.1, this ball is isomorphic with a ball of the planar graph $G(m)$. But this ball is itself planar and in a planar graph, every cycle is a boundary. Thus the group $H_1(G_\varepsilon)$ is trivial. □

The next lemma will allow us to compute the dimension of the induced homology group of every subgraph G_ε of $G = G_t(m)$. Since a set $\varepsilon \subset E$ can be regarded as a subset of E^*, it also defines a subgraph G_ε^* of the graph G^*. Let us denote by $\operatorname{rank} G_\varepsilon$ ($\operatorname{rank} G_\varepsilon^*$) the rank of an incidence matrix of G_ε (of G_ε^*). By Lemma 3.2 these ranks do not depend on the choice of the incidence matrix of the graph. The dimension of the induced homology group is given by:

Lemma 4.3. *For every $\varepsilon \subset E$, we have*

$$\dim H_1(G_\varepsilon) = |\varepsilon| - \frac{2}{m}|E| + 1 + \operatorname{rank} G_\varepsilon^* - \operatorname{rank} G_\varepsilon.$$

Proof. The group $H_1(G_\varepsilon)$ is the quotient of the cycle code of G_ε by $\operatorname{Im} \partial_2^\varepsilon$, the set of boundaries of G which are included in the subgraph ε.

By definition, the cycle code of G_ε is the kernel of the map ∂_1^ε. Moreover, the incidence matrix of G_ε is a matrix of this linear map. Therefore, the dimension of the cycle code of the subgraph G_ε is

$$\dim \ker \partial_1^\varepsilon = |\varepsilon| - \operatorname{rank} G_\varepsilon. \tag{3}$$

The set of boundaries of G is the image of the map ∂_2. We noticed in Section 3.1 that a matrix of the map ∂_2 is given by the transpose of $B(G^*)$, the incidence matrix of G^*. This means that the boundaries of G correspond to the sums of rows of $B(G^*)$. These are the vectors of the form $xB(G^*)$, where x is a binary vector.

Consider the incidence matrix of $G^*_{\bar{\varepsilon}}$, where $\bar{\varepsilon}$ denotes the complement of ε in E. This matrix can be obtained from $B(G^*)$ by selecting the columns indexed by the edges in $\bar{\varepsilon}$. Let us define a map ϕ which sends a sum of rows of $B(G^*)$ onto the same sum of rows in the matrix $B(G^*_{\bar{\varepsilon}})$. It is the map

$$\begin{aligned} \phi : \operatorname{Im} \partial_2 &\longrightarrow C_1^{\bar{\varepsilon}} \\ xB(G^*) &\longmapsto x_{\bar{\varepsilon}} B(G^*_{\bar{\varepsilon}}), \end{aligned}$$

where x is a row vector of $\mathbb{F}_2^{|V|}$ and $x_{\bar{\varepsilon}}$ is its restriction to the columns indexed by the edges of $\bar{\varepsilon}$. Then, the boundaries of G included in ε, are exactly the vectors of the kernel of ϕ. The dimension of this space is

$$\dim \operatorname{Im} \partial_2^{\varepsilon} = \dim \ker \phi = \dim \operatorname{Im} \partial_2 - \dim \operatorname{Im} \phi = \operatorname{rank} G^* - \operatorname{rank} G^*_{\bar{\varepsilon}}. \tag{4}$$

Now $\operatorname{rank} G^* = \dim \operatorname{Im} \partial_1^* = |E^*| - \dim \ker \partial_1^*$. Applying Lemma 3.2 to the dimension of the cycle code $\ker \partial_1^*$ of G^* and the fact that $G = G_t(m)$ is connected, we get $\operatorname{rank} G^* = |F| - 1 = (2/m)|E| - 1$. Injecting this last fact into (4), we obtain, together with (3), the formula for $\dim H_1(G_\varepsilon) = \dim \ker \partial_1^\varepsilon - \dim \operatorname{Im} \partial_2^\varepsilon$. □

4.2 The rank difference function

We now consider the probabilistic behaviour of the induced homology of a random subgraph of $G_t = G_t(m)$. To get a distribution which locally coincides with the distribution of percolation events, the subset of edges ε is chosen by selecting each edge of G_t independently with probability p. This defines a random subgraph $G_{t,\varepsilon}$ of the graph G_t.

The intuition we follow is that if we are below the critical probability of the graph $G(m)$, then most connected components appearing in the random subgraph $G_{t,\varepsilon}$ should be small. Thanks to Lemma 4.2, these clusters do not support any non trivial homology. This implies that if $p < p_c(G(m))$ then the dimension of the induced homology of $G_{t,\varepsilon}$ must be small. Conversely, if we compute, using Lemma 4.3, the expected dimension of $H_1(G_{t,\varepsilon})$ and find it to be large, we know that p must be above the critical probability p_c. These considerations lead us to introduce the following quantity.

Definition 4.4. *The rank difference function associated with the family of graphs $(G_t)_t$ is defined to be*

$$D(p) = \limsup_t \mathbb{E}_p \left(\frac{\operatorname{rank} G^*_{t,\bar{\varepsilon}} - \operatorname{rank} G_{t,\varepsilon}}{|E_t|} \right).$$

The rank difference function satisfies the folowing equation when p is below the critical probability of $G(m)$.

Theorem 4.5. *If $p < p_c(G(m))$ then the rank difference function associated with the family $(G_t)_t$ satifies*

$$p - \frac{2}{m} + D(p) = 0.$$

Corollary 4.6. *Defining $p_h = \sup\{p, p - \frac{2}{m} + D(p) = 0\}$ we have $p_c \leq p_h$.*

Assume that $p < p_c(G(m))$. By definition of the critical probability, for any fixed edge e of the infinite graph $G(m)$, the probability that e is contained in an open connected component $C(e)$ of $G(m)$ of size strictly larger than t vanishes when $t \to \infty$. The following lemma shows that we observe a similar behaviour in the finite graphs G_t. It will be instrumental in proving Theorem 4.5.

Lemma 4.7. *For every $t \geq 0$, fix an edge e_t of the graph G_t and denote by $C(e_t)$ its (possibly empty) connected component in the random subgraph $G_{t,\varepsilon}$. Then, the probability that $C(e_t)$ contains strictly more than $t-2$ edges tends to 0 when t goes to infinity.*

Proof. The complementary event depends only on what occurs inside the ball of radius t centered on an endpoint of the edge e_t. Since this ball is isomorphic to the ball with the same radius in $G(m)$, this event has the same probability in the space $G(m)$ and in $G_t(m)$. Hence the result by the remark preceding the lemma. □

Proof of Theorem 4.5.

Thanks to Lemma 4.1, we have the following decomposition of the first homology group of $G_{t,\varepsilon}$:

$$H_1(G_{t,\varepsilon}) = \bigoplus_{i=1}^{r} H_1(G_{t,\varepsilon_i}).$$

where ε_i is the edge set of the i-th connected component of $G_{t,\varepsilon}$.

From Lemma 4.2, all the components ε_i of size smaller than t have a trivial contribution to $H_1(G_{t,\varepsilon})$. For the other components, the dimension of $H_1(G_{t,\varepsilon_i})$ is bounded by the number of edges in the component ε_i. Indeed, the induced homology group of G_{t,ε_i} is a quotient of the cycle code of this graph, whose dimension is at most the number of egdes in ε_i. This implies

$$\dim H_1(G_{t,\varepsilon}) \leq |\{e \in E_t \text{ such that } |C(e)| > t\}|,$$

where $C(e)$ denotes the connected component in $G_{t,\varepsilon}$ of the edge e and $|C(e)|$ is its number of edges.

Let us denote by $X_t = X_t(G_{t,\varepsilon})$ the cardinality of the set $\{e \in E_t \text{ such that } |C(e)| > t\}$. To study the expectation of X_t, we define a random variable X_e, associated with each edge $e \in E_t$, which takes the value $X_e(G_{t,\varepsilon}) = 1$ if the size of $C(e)$ is larger than t and which is 0 otherwise. Consequently, we have

$$X_t = \sum_{e \in E_t} X_e,$$

and by linearity of expectation, $\mathbb{E}(X_t) = \sum_e \mathbb{E}(X_e)$. For every edge $e \in E_t$, this expectation of the random variable X_e is $\mathbb{E}(X_e) = \mathbb{P}(|C(e)| > t)$. By edge-transitivity of the graph G_t, this quantity does not depend on the edge e, thus $\mathbb{E}(X_t) = |E_t| \, \mathbb{P}(|C(e_t)| > t)$, for some fixed edge e_t of the graph G_t. Moreover, from Lemma 4.7, this probability vanishes when t goes to infinity. This allows us to bound the expected dimension of the induced homology:

$$\mathbb{E}_p\left(\frac{\dim H_1(G_{t,\varepsilon})}{|E_t|}\right) \leq \mathbb{E}_p\left(\frac{X_t}{|E_t|}\right) = \mathbb{P}_p(|C(e_t)| > t) \to 0.$$

Since the right-hand side tends to 0 when t goes to infinity, taking the superior limit gives exactly 0, *i.e.*

$$\limsup_t \mathbb{E}_p\left(\frac{\dim H_1(G_{t,\varepsilon})}{|E_t|}\right) = 0.$$

To conlude the proof, we determine the expected dimension of the induced homology group with the help of Lemma 4.3. We find

$$\limsup_t \mathbb{E}_p \left(\frac{\dim H_1(G_{t,\varepsilon})}{|E_t|} \right) = p - \frac{2}{m} + D(p).$$

□

5 Computation of the rank difference function of hyperbolic tilings

The behaviour of the function $D(p)$ is difficult to capture directly from its definition. The aim of this section is to provide an explicit combinatorial description of the rank difference function $D(p)$ associated with the finite tilings $(G_t)_t$.

The next lemma enables us to replace the rank which appears in the definition of $D(p)$ by a strictly graph-theoretical quantity.

Lemma 5.1. *Let $\kappa_{t,\varepsilon}$ denote the number of connected components of the graph $G_{t,\varepsilon}$. We have:*

$$\operatorname{rank} G_{t,\varepsilon} = |V_t| - \kappa_{t,\varepsilon}.$$

Proof. By definition, the rank of the graph $G_{t,\varepsilon}$ is the rank of an incidence matrix of this graph. The kernel of this incidence matrix is the cycle code of the graph $G_{t,\varepsilon}$, which has dimension $|\varepsilon| - |V_t| + \kappa_{t,\varepsilon}$ from Lemma 3.2. The result follows. □

The function $D(p)$ depends on the expected rank of the random submatrix $G_{t,\varepsilon}$. This encourages us to examine the expected number of connected components of the random subgraph $G_{t,\varepsilon}$. A key ingredient of our study is the following decomposition of the random variable $\kappa_{t,\varepsilon}$.

Lemma 5.2. *Let C be a connected subgraph of G_t. Denote by X_C the random variable which takes the value 1 if C is a connected component of the random graph $G_{t,\varepsilon}$ and 0 otherwise. Then, we have*

$$\kappa_{t,\varepsilon} = \sum_{C \in \mathcal{C}_t} X_C$$

where $\mathcal{C}_t$ denotes the set of connected subgraphs C of $G_t(m)$.

Moreover, we have $\mathbb{E}_p(X_C) = p^{|E(C)|}(1-p)^{|\partial(C)|}$ where $\partial(C)$ is the set of edges of G_t which are incident to at least one vertex of C, but which do not belong to $E(C)$.

The proof of the above lemma is self-evident. Using this decomposition of $\kappa_{t,\varepsilon}$, we derive the following exact expression of the rank difference function as a function of the subgraphs of the infinite graph $G(m)$.

Theorem 5.3. *For $m \geq 5$ and $0 < p \leq 1/2$, The rank difference function associated with the graphs $(G_t(m))_t$ is equal to*

$$D(p) = \frac{2}{m} \sum_{C \in \mathcal{C}(v)} \left(\frac{1}{|V(C)|} \left(p^{|E(C)|}(1-p)^{|\partial(C)|} - (1-p)^{|E(C)|} p^{|\partial(C)|} \right) \right),$$

where $\mathcal{C}(v)$ denotes the set of connected subgraphs C of $G(m)$ containing a fixed vertex v.

Proof. From Lemma 5.1, the rank difference function can be rewritten

$$\begin{aligned} D(p) &= \limsup_t \mathbb{E}_p\left(\frac{\kappa_{t,\varepsilon}-\kappa_{t,\bar\varepsilon}}{|E_t|}\right) \\ &= \limsup_t \left(\mathbb{E}_p\left(\frac{\kappa_{t,\varepsilon}}{|E_t|}\right) - \mathbb{E}_{1-p}\left(\frac{\kappa_{t,\varepsilon}}{|E_t|}\right)\right). \end{aligned}$$

where we used the fact that, $\bar\varepsilon$ being the complement of ε in E_t, we have $\mathbb{E}_p(\kappa_{t,\bar\varepsilon}) = \mathbb{E}_{1-p}(\kappa_{t,\varepsilon})$.

Then, using the decomposition of $\kappa_{t,\varepsilon}$ proposed in Lemma 5.2 and the linearity of expectation, we obtain

$$D(p) = \limsup_t \frac{1}{|E_t|} \sum_{C\in\mathcal{C}_t} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right).$$

Elimination of the large components— Now, remark that the main contribution in this sum is given by the small components. To prove this, consider a sequence of integers $(M_t)_t$ such that $M_t \to +\infty$. Then, we have

$$\begin{aligned} \frac{1}{|E_t|}\sum_{\substack{C\in\mathcal{C}_t\\ |E(C)|\geq M_t}} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right) &\leq \frac{1}{|E_t|}\sum_{\substack{C\in\mathcal{C}_t\\ |E(C)|\geq M_t}} \left(\mathbb{E}_p(X_C) + \mathbb{E}_{1-p}(X_C)\right) \\ &= \frac{1}{|E_t|}\mathbb{E}_p\left(\sum_{\substack{C\in\mathcal{C}_t\\ |E(C)|\geq M_t}} X_C\right) + \frac{1}{|E_t|}\mathbb{E}_{1-p}\left(\sum_{\substack{C\in\mathcal{C}_t\\ |E(C)|\geq M_t}} X_C\right) \\ &\leq \frac{1}{|E_t|}\frac{2|E_t|}{M_t} = \frac{2}{M_t} \to 0 \end{aligned}$$

To obtain the last inequality, remark that the sum of all the random variables X_C such that $|E(C)| \geq M_t$ counts the number of connected components of the subgraph $G_{t,\varepsilon}$ of size larger than M_t. Since connected components are disjoint, this number cannot be larger than $|E_t|/M_t$.

The previous paragraph proves that, for every sequence M_t going to infinity, the rank difference function is given by

$$D(p) = \limsup_t \frac{1}{|E_t|}\sum_{\substack{C\in\mathcal{C}_t\\ |E(C)|< M_t}} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right)$$

Recentralization— In order to remove the dependency on t, we would like to apply the local isomorphism between $G_t(m)$ and $G(m)$ and to express everything as a function of the infinite graph $G(m)$. First, we have to recenter all the components C around a fixed vertex v_t of the graph G_t. To move a connected component C of the graph G_t onto a component which contains the vertex $v = v_t$, we use a family of automorphisms of the graph $G_t(m)$. For every vertex w of the graph $G_t(m)$, select $\sigma_{v,w}$, an automorphism of the graph $G_t(m)$ sending v onto w. We take the identity for $\sigma_{v,v}$. Such an automorphism exists because the graph G_t is vertex transitive, as explained in Section 2. From this fixed family of automorphisms, we can reach all the connected subgraphs of G_t, starting from the subgraphs containing v. Stated differently, we have

$$\mathcal{C}_t = \{C \mid C \text{ connected } \} = \bigcup_{w \in V_t} \{\sigma_{v,w}(C) \mid C \text{ connected }, v \in V(C)\}$$

At the right-hand side of this equality, each component C of the graph appears $|V(C)|$ times. Moreover, the contribution $\mathbb{E}_p(X_C)$ of the subgraph C, computed in Lemma 5.2, depends only on $|E(C)|$ and $|\partial(C)|$, which are both invariant under the application of an automorphism $\sigma_{v,w}$. Hence, $D(p)$ is equal to

$$\begin{aligned} D(p) &= \limsup_t \frac{1}{|E_t|} \sum_{\substack{C \in \mathcal{C}_t \\ |E(C)| < M_t}} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right) \\ &= \limsup_t \frac{1}{|E_t|} \sum_{\substack{C \in \mathcal{C}_t(v) \\ |E(C)| < M_t}} \sum_{w \in V_t} \frac{1}{|V(C)|} \left(\mathbb{E}_p(X_{\sigma_{v,w}(C)}) - \mathbb{E}_{1-p}(X_{\sigma_{v,w}(C)})\right) \\ &= \limsup_t \frac{1}{|E_t|} \sum_{\substack{C \in \mathcal{C}_t(v) \\ |E(C)| < M_t}} \frac{|V_t|}{|V(C)|} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right) \\ &= \limsup_t \frac{2}{m} \sum_{\substack{C \in \mathcal{C}_t(v) \\ |E(C)| < M_t}} \frac{1}{|V(C)|} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right) \end{aligned}$$

where we have used $\frac{|V_t|}{|E_t|} = \frac{2}{m}$ since G_t is m-regular.

Application of the local isomorphism— We now replace the graph $G_t(m)$ by the infinite graph $G(m)$. Since the balls of radius t are isomorphic in $G_t(m)$ and in $G(m)$, we have that every fixed subgraph C inside such a ball has the same probability of being a connected component whether it is of the random subgraph $G_{t,\varepsilon}$ or of the open subgraph of $G(m)$. By choosing $M_t = t - 1$, we therefore get

$$D(p) = \limsup_t \frac{2}{m} \sum_{\substack{C \in \mathcal{C}(v) \\ |E(C)| < M_t}} \frac{1}{|V(C)|} \left(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C)\right) \tag{5}$$

where $\mathcal{C}(v)$ denotes the set of connected subgraphs C of $G(m)$ containing the fixed vertex v.

We can now conclude the proof. From Lemma 5.2, the quantity $(\mathbb{E}_p(X_C) - \mathbb{E}_{1-p}(X_C))$ is equal to $\left(p^{|E(C)|}(1-p)^{|\partial(C)|} - (1-p)^{|E(C)|}p^{|\partial(C)|}\right)$, which is positive by Lemma 5.4 to be proven just below. Therefore all the terms of the sum in (5) are positive, which means that the lim sup is in fact a limit. Since $M_t \to +\infty$, we get

$$D(p) = \frac{2}{m} \sum_{C \in \mathcal{C}(v)} \left(\frac{1}{|V(C)|} \left(p^{|E(C)|}(1-p)^{|\partial(C)|} - (1-p)^{|E(C)|}p^{|\partial(C)|} \right) \right).$$

□

It remains to prove that the series has positive terms. This result relies on an isoperimetric inequality.

Lemma 5.4. *Let $0 < p < 1/2$. For every connected subgraph C of $G(m)$, we have*

$$p^{|E(C)|}(1-p)^{|\partial(C)|} - (1-p)^{|E(C)|}p^{|\partial(C)|} > 0.$$

Proof. The parameter p is assumed to be smaller than $1/2$. Thus, to prove that this quantity is strictly positive it suffices to show that for every connected subgraph C of $G(m)$, we have $|E(C)| < |\partial(C)|$. This inequality is somewhat analogous to the isoperimetric inequality that we recall now. The isoperimetric constant of the graph $G(m)$ is defined to be

$$i_E(G(m)) = \inf \left\{ \frac{|\partial(C)|}{|V(C)|} \right\}$$

with C ranging over all finite subgraphs (that can be assumed connected) of $G(m)$. This number was computed exactly for hyperbolic graphs in [11]. It is

$$i_E(G(m)) = (m-2)\sqrt{1 - \frac{4}{(m-2)^2}}. \tag{6}$$

In order to apply this to our problem, we write

$$\frac{|\partial(C)|}{|E(C)|} = \frac{|\partial(C)|}{(m/2)|V(C)| - (1/2)|\partial(C)|} \geq \frac{i_E(G(m))}{m/2 - i_E(G(m))/2} \tag{7}$$

where we have used the fact that the smallest rate $|\partial(C)|/|E(C)|$ is achieved when $\partial(C)$ contains only edges with exactly one endpoint in C. In that case, we have $m|V(C)| = 2|E(C)| + |\partial(C)|$. Using Equation (6) and (7), it is then easy to check that, for all $m \geq 5$, we have

$$\frac{|\partial(C)|}{|E(C)|} \geq \frac{i_E(G(5))}{5/2 - i_E(G(5))/2} \approx 1.62 > 1.$$

This proves the lemma. □

6 Bound on the critical probability of the hyperbolic lattice $G(m)$

We showed in Theorem 4.5 that the critical probability of $G(m)$ is bounded from above as $p_c(G(m)) \leq p_h$ with p_h defined in Corollary 4.6. Theorem 5.3 provides an exact formula for the rank difference function $D(p)$ as a sum of a series depending on the connected subgraphs of $G(m)$. This gives a new expression for p_h that does not involve the finite graphs $G_t(m)$ anymore, but it still leaves p_h difficult to compute. We now show that by replacing the series $D(p)$ by its partial sums, we obtain explicit upper bounds on p_h and hence on p_c.

Theorem 6.1. *Let $n \geq 0$ and let $D_n(p)$ be a partial sum of the series $D(p)$ associated with the hyperbolic graph $G(m)$. Then, the solution $p_h(n) \in [0,1]$ of the equation*

$$p - 2/m + D_n(p) = 0$$

is an upper bound on p_h and hence on $p_c(G(m))$.

Proof. We have seen in Lemma 5.4 that all the terms of the series $D(p)$ are strictly positive when $p > 0$. Thus, every partial sum $D_n(p)$ satisfies $D_n(p) < D(p)$. As a consequence, if $p_h(n)$ is a solution of the equation $p - 2/m + D_n(p) = 0$, then we have $p_h(n) - 2/m + D(p_h(n)) > 0$. This proves that $D(p)$ does not satisfy the criterion of Theorem 4.5 at $p = p_h(n)$. Therefore $p_h(n)$ is an upper bound on p_h. □

As a first application of this theorem, using only the fact that $D_n(p) \geq 0$, we recover the upper bound $p_c(G(m)) \leq 2/m$, proved in [6].

The first terms of the series, corresponding to the components of small size can be computed easily. For example the number of connected subgraphs of size 0, that is with 0 edges, containing a fixed vertex of $G(m)$ is 1 and this subgraph has a boundary $\partial(C)$ of size m. This gives the partial sum

$$D_0(p) = \frac{2}{m}((1-p)^m - p^m).$$

Applying Theorem 6.1 to $D_0(p)$, we get an upper close to 0.35. This is already more precise than the upper bound in [7].

The next partial sum is given by

$$D_1(p) = D_0(p) + \frac{2}{m}\left(\frac{m}{2}(p(1-p)^{2(m-1)} - p^{2(m-1)}(1-p))\right),$$

since there are m different connected subgraphs of $G(m)$ composed of one edge and containing a fixed vertex.

The first terms can be computed easily in this way. In a tree it is possible to get an exact formula for the number of rooted connected subgraphs using the Lagrange inversion threorem. However this enumeration problem becomes extremely difficult when the subgraphs start covering cycles. Moreover, the size of the boundary and the number of vertices of the subgraph do not depend only on its number of edges. We enumerated all the connected subgraphs of $G(5)$ of size at most 8 by computer. The results are given in Table 1. Using the partial sum $D_8(p)$ that takes into acount all the subgraphs of size at most 8, we get an upper bound on $p_c(G(5))$ which is approximately 0.300007:

$$p_c(G(5)) \leq 0.300007.$$

To the best of our knowledge, the previous best upper bound was close to 0.38 [7].

7 Concluding comments

Summarising Theorems 4.5 and 5.3 we have proved :

Theorem 7.1. *For $m \geq 5$ we have $p_c(G(m)) \leq p_h$ with*

$$p_h = \sup\{p \in [0, 1/2] \mid D(p) + p - \frac{2}{m} = 0\} \text{ and}$$
$$D(p) = \frac{2}{m} \sum_{C \in \mathcal{C}(v)} \left(\frac{1}{|V(C)|}\left(p^{|E(C)|}(1-p)^{\partial(C)} - (1-p)^{|E(C)|}p^{\partial(C)}\right)\right)$$

where $\mathcal{C}(v)$ denotes the set of connected subgraphs C of $G(m)$ containing a fixed vertex v of the graph $G(m)$.

Table 1: Enumeration of the rooted subgraphs of $G(5)$ up to size 8.

$\|E(C)\|$	$\|V(C)\|$	$\partial(C)$	occurence
0	1	5	1
1	2	8	5
2	3	11	30
3	4	14	200
4	5	17	1400
4	5	16	25
5	6	20	10146
5	6	19	450
5	5	15	5
6	7	23	75460
6	7	22	5775
6	6	18	90
7	8	26	572720
7	8	25	64200
7	8	24	480
7	7	21	1155
8	9	29	4418190
8	9	28	661950
8	9	27	13005
8	8	24	12840
8	8	23	180

The value p_h can be thought of as a critical value for the appearance of homology in the graph $G(m)$. It captures the following threshold : for $p > p_h$, open subgraphs of large finite versions of $G(m)$ must have a first homology group of dimension that scales linearly with the total number of edges of the finite graph. For $p < p_h$, the dimension of the homology group is sublinear instead. This bound is really meaningful only for the hyperbolic case $m \geq 5$ since for $m = 4$ (the square lattice), the dimension of the total homology group of finite versions of the infinite grid (tori) is limited to 2.

A consequence of Theorem 7.1 is that p_h gives an upper bound on the parameters of the quantum erasure channel that hyperbolic surface codes built on the family $G_t(m)$ can sustain [7].

We conjecture :

Conjecture 7.2. *For $m \geq 5$, $p_c = p_h$.*

Recall that in hyperbolic lattices it has been shown that immediately beyond the critical probability, the open subgraph contains infinitely many infinite connected components [3]. The conjecture could be seen as a "finite" (but unbounded) version of this fact.

Acknowledgements

Nicolas Delfosse was supported by the Lockheed Martin Corporation.

References

[1] S. K. Baek, P. Minnhagen, and B. Jun Kim. Percolation on hyperbolic lattices. *Phys. Rev. E*, 79:011124, Jan 2009.

[2] I. Benjamini, A. Nachmias and Y. Peres. Is the critical percolation probability local ? *Probability Theory and Related Fields*, 149, pp. 261–269, 2011.

[3] I. Benjamini and O. Schramm. Percolation beyond Z^d, many questions and a few answers. *Electr. Commun. Probab*, 1:71–82, 1996.

[4] I. Benjamini, O. Schramm. Percolation in the hyperbolic plane. *J. Am. Math. Soc.*, 29 pp. 487–507, 2001.

[5] C. Berge. *Graphs and Hypergraphs.* Elsevier, 1973.

[6] N. Delfosse and G. Zémor. Quantum erasure-correcting codes and percolation on regular tilings of the hyperbolic plane. In *Proc. of IEEE Information Theory Workshop, ITW 2010*, pages 1–5, 2010.

[7] N. Delfosse and G. Zémor. Upper bounds on the rate of low density stabilizer codes for the quantum erasure channel. *Quantum Information & Computation*, 13(9-10):793–826, 2013.

[8] P. Giblin. *Graphs, surfaces and homology.* Cambridge University Press, 2010.

[9] G. Grimmett. *Percolation.* Springer-Verlag, New York, 1989.

[10] H. Gu and R. M. Ziff. Crossing on hyperbolic lattices. *Phys. Rev. E* 85:051141, May 2012.

[11] O. Haggstrom, J. Jonasson, and R. Lyons. Explicit isoperimetric constants and phase transitions in the random-cluster model. *Annals of Probability*, pages 443–473, 2002.

[12] A. Hatcher. *Algebraic topology.* Cambridge University Press, 2002.

[13] H. Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Communications in Mathematical Physics*, 74:41–59, 1980.

[14] J.F. Lee and S.K. Baek. Bounds of percolation thresholds on hyperbolic lattices. *Phys. Rev. E* 86, 062105 2012.

[15] J. Širáň. Triangle group representations and constructions of regular maps. *Proc. of the London Mathematical Society*, 82(03):513–532, 2000.