Determining the Genus of a Map by Local Observation of a Simple Random Process

Itai Benjamini and László Lovász
Microsoft Research
One Microsoft Way, Redmond, WA 98052
e-mail: itai@wisdom.weizmann.ac.il lovasz@microsoft.com

Abstract

Given a graph embedded in an orientable surface, a process with stationary increments, consisting of random excitations and random node and face balancing, is constructed and analyzed. It is shown that, given a priori bounds $\overline{g}$ on the genus and $\overline{n}$ on the number of nodes, one can determine the genus of the surface from local observations of the process, restricted to any connected subgraph which cannot be separated from the rest of the graph by fewer than $16\overline{g}$ nodes. The observation time and the computation time are polynomial in $\overline{n}^{\overline{g}}$.

The process constructs slightly perturbed random "discrete analytic functions" on the surface, and the key fact in the analysis is that such a function cannot vanish on a large piece of the surface.

1 Introduction

At least since Polya's observation (1921) that "a drank man will return home while a drank bird might lose its way forever", it is known that geometric properties of the underling space can manifest themselves in the behavior of a random process taking place on the space. Moreover, in recent years random processes were used to retrieve information about the underlying space. E.g. sampling, volume estimates of convex bodies or scenery reconstruction along random walk paths (see for instance [9], [2] and [8]). Rather then letting a random walker wonder around and gather information, in this note we would like to study a problem in which a random process is observed locally in a fixed bounded neighborhood and still non-trivial global observation can be distilled via these observations.

In addition to the algorithmic motivation, a reasonable question might be: Consider a stationary spin system on a graph, such as Glauber dynamics for the Ising model on a graph. What properties of the graph can be inferred from properties of the process (for instance, are there interesting relations between the spectrum of the Ising dynamics and the graph spectrum)? A harder challenge is to infer non-trivial global properties of the underling graph from local observations of the process.

Facing the harder challenge one might devise first custom made variants of the standard spin systems, which can be analyzed and provide ways to compute global invariants using "physical" systems.

Indeed, below we will study a reasonably natural and simple process, called *noisy circulator*, with local operations, living on the edges of a graph, embedded in an orientable surface. The noisy circulator consists of adding mass 1 to randomly chosen edges with a slow rate and balancing the flow into vertices or around faces at random and with a faster rate (details below). Somewhat surprisingly, it will be shown how to extract, almost surely and in polynomial time in the size of the graph, the genus of the surface by observing the restriction of the process to a bounded set of edges, depending only on an a priori bound on the genus.

Although from a pure algorithmic view point, this first construction might be useful and might have some advantages, the point is not to devise the optimal ad-hoc distributed algorithm for finding the genus, under certain restrictions, but to show how observing locally a simple physical process already does that, and to present this computation scheme. It is of interest then to construct other examples of similar flavor.

The structure of the proof is twofold, a topological theorem and a statistical element.

In the topological part, we define *smooth circulations*: circulations that are also circulations on the dual map. These can be considered as discrete analogues of analytic functions. Every homology class of circulations contains exactly one smooth circulation, so the dimension of their space can be used to find the genus of the surface. A key result (theorem 7), which is of independent interest, asserts that every connected piece of the vanishing set of a smooth circulations can be separated from the rest of the graph by a small number of nodes.

In the statistical part, we determine the dimension of the space of smooth circulations, from observations which can be considered as samples of smooth circulations, restricted to a bounded set of edges, with an additional noise.

The next subsection contains some necessary definitions and basic propositions, in subsection 1.2 a description of the process and a formulation of the main theorem are given, together with an outline of the proof. The rest of the paper contains the proof of the main theorem. In the final section we end with some further comments and problems.

1.1 Circulations and homology

Let S be a closed compact surface, and consider a *map* on S, i.e., a graph $G = (V, E)$ embedded in S so that each face is a disc. We can describe the map as a triple $G = (V, E, \mathcal{F})$, where V is the set of nodes, E is the set of edges, and $\mathcal{F}$ is the set of faces of G. We fix a reference orientation of G; then each edge $e \in E$ has a *tail* $t(e) \in V$, a *head* $h(e) \in V$, a *right shore* $r(e) \in \mathcal{F}$, and a *left shore* $l(e) \in \mathcal{F}$.

The embedding of G defines a *dual map* G^*. Combinatorially, we can think of G^* as the triple $(\mathcal{F}, E, V)$, where the meaning of "node" and "face", "head" and "right shore", and "tail" and "left shore" is interchanged. (Taking the dual of the dual will give the original map with every edge reversed; this should not concern us in this paper.)

For each edge e, let $\chi_e \in \mathbb{R}^E$ be the unit vector that is 1 on e and 0 elsewhere; we define $\chi_v \in \mathbb{R}^V$ for $v \in V$ and $\chi_F \in \mathbb{R}^{\mathcal{F}}$ for $F \in \mathcal{F}$ analogously.

For each node v, let $\delta v \in \mathbb{R}^E$ denote the coboundary of v:

$$(\delta v)_e = \begin{cases} 1, & \text{if } h(e) = v, \\ -1, & \text{if } t(e) = v, \\ 0, & \text{otherwise.} \end{cases}$$

Thus $|\delta v|^2 = d_v$ is the degree of v. A vector $\phi \in \mathbb{R}^E$ is a *circulation* if

$$\phi \cdot \delta v = \sum_{e:\ h(e)=v} \phi(e) - \sum_{e:\ t(e)=v} \phi(e) = 0.$$

For every face $F \in \mathcal{F}$, we denote by $\partial F \in \mathbb{R}^E$ the boundary of F:

$$(\partial F)_e = \begin{cases} 1, & \text{if } r(e) = F, \\ -1, & \text{if } l(e) = F, \\ 0, & \text{otherwise.} \end{cases}$$

Then $d_F = |\partial F|^2$ is the length of the cycle bounding F.

Each vector ∂F is a circulation; circulations that are linear combinations of vectors ∂F are called *null-homologous*. Two circulations ϕ and ϕ' are called *homologous* if $\phi - \phi'$ is null-homologous.

Let ϕ be a circulation on G. We say that ϕ is *smooth* if for every face $F \in \mathcal{F}$, we have

$$\phi \cdot \partial F = 0.$$

This is equivalent to saying that ϕ is a circulation on the dual map G^*.

Remark 1 *Smooth circulations are closely related to discrete analytic functions and are essentially the same as discrete holomorphic 1-forms. These functions were introduced for the case of the square grid a long time ago [6, 5]. For the case of a general planar graph, the notion is implicit in [3]. For a detailed treatment see [10].*

To explain the connection, let ϕ be a smooth circulation on a graph G embedded in a surface. Consider a planar piece of the surface. Then on the set $\mathcal{F}'$ of faces contained in this planar piece, we have a function $\sigma: \mathcal{F}' \to \mathbb{R}$ such that $\partial\sigma = \phi$, i.e., $\phi(e) = \sigma(r(e)) - \sigma(l(e))$ for every edge e. Similarly, we have a function $\pi: V' \to \mathbb{R}$ (where V' is the set of nodes in this planar piece), such that $\delta\pi = \phi$, i.e., $\phi(e) = \pi(t(e)) - \pi(h(e))$ for every edge e. We can think of π and σ as the real and imaginary parts of a (discrete) analytic function. The relation $\delta\pi = \rho\phi$ is then a discrete analogue of the Cauchy–Riemann equations.

Thus we have the two linear orthogonal subspaces $\mathcal{A} \subseteq \mathbb{R}^E$ generated by the vectors δv $(v \in V)$ and $\mathcal{B} \subseteq \mathbb{R}^E$ generated by the vectors ∂F $(F \in \mathcal{F})$. Vectors in $\mathcal{B}$ are 0-homologous circulations. The orthogonal complement $\mathcal{A}^\perp$ is the space of all circulations, and $\mathcal{B}^\perp$ is the space of circulations on the dual graph. The intersection $\mathcal{C} = \mathcal{A}^\perp \cap \mathcal{B}^\perp$, the space of smooth circulations. So $\mathbb{R}^E = \mathcal{A} \oplus \mathcal{B} \oplus \mathcal{C}$. From this picture we conclude the following.

Proposition 2 *Every circulation is homologous to a unique smooth circulation.*

It also follows that $\mathcal{C}$ is isomorphic to the first homology group of S (over the reals), and hence we get the following:

Proposition 3 *The dimension of the space $\mathcal{C}$ of smooth circulations is $2g$.*

1.2 Main result: randomized circulations

1.2.1 The Noisy Circulator.

We consider the following process on G. Let $p > 0$ be fixed. We start with the vector $x = 0 \in \mathbb{R}^E$. At each further step, the following two operations are carried out on the current vector $x \in \mathbb{R}^E$:

(a) [Node balancing.] We choose a random node v, and subtract from x the vector $(x^\mathsf{T}\delta_v/d_v)\delta_v$.

(b) [Face balancing.] We choose a random face F, and subtract from x the vector $(x^{\mathsf{T}}\partial_F/d_F)\partial_F$.

In addition, with some given probability $p > 0$, we do the following:

(c) [Excitation.] We choose a random edge e, and add 1 to x_e.

Immediately after a node balancing step, the node v just balanced satisfies the flow condition; a subsequent other node balancing, or a face balancing may destroy this. Smooth circulations are invariant under node and face balancing, and we'll see that under repeated application of (a) and (b), any vector converges to a smooth circulation.

1.2.2 Observing the process.

Let $U \subseteq V(G)$ induce a connected subgraph of G, and assume that every cycle that separates S into two parts S_1 and S_2 so that S_1 contains U and S_2 is not a disc, has length at least $16g$. We show that if p is small enough, then observing the process on the edges incident with U long enough, we can determine the genus of the surface.

To be precise, we do the following. Let E_0 be the set of edges incident with U. Let $x(t) \in \mathbb{R}^E$ be the vector after t steps, and let $y(t)$ be the restriction of $x(t) \in \mathbb{R}^{E_0}$ to the edges in E_0. So we can observe the sequence random vectors $y(0), y(1), \dots$.

Let $x'(t)$ be the projection of $x(t)$ onto $\mathcal{C}$, and let $y'(t)$ be the restriction of $x'(t)$ to the edges in E_0. Because of the random steps (c), after a sufficiently long time, the vectors $x'(0), \dots, x'(t-1)$ will span the space $\mathcal{C}$. So the rank of this set of vectors gives us the dimension of $\mathcal{C}$, and so by Proposition 3, it gives us the genus of the surface. The restriction to E_0 is one-to-one on $\mathcal{C}$ (Theorem 7), and so this rank is the same as the rank of the set $\{y'(0), y'(1), \dots, y'(t)\}$.

Unfortunately, we cannot observe the vectors $y'(t)$, only the vectors $y(t)$. But (at least if p is small) we expect $y(t)$ to be close to $y'(t)$. Indeed, the "errors" $x''(t) = x(t) - x'(t)$ are in $\mathcal{A} \oplus \mathcal{B}$, and therefore it is not hard to show that they tend to 0 exponentially fast (see lemma 10 below), at least as long as no "excitation" step (c) occurs. Therefore the restrictions $y''(t) = y(t) - y'(t)$ also tend to 0 exponentially. The speed of convergence depends on the eigenvalue gap of the transition matrix of the (undirected) random walk on G and G^*.

So we are lead to the following standard statistical problem: there is a sequence $y'(0), y'(1), \dots$ of vectors in $\mathbb{R}^k$, which span a linear subspace L. We observe the sequence $y(t) = y'(t) + y''(t)$, where the "error" $y''(t)$ is small on the average. We want to find the dimension k.

The random vectors $x(t)$ or $x'(t)$ are not independent; but if we take the differences, i.e., we look at the vectors $x'(t+1) - x'(t)$, then these are independent (in the probabilistic sense). Indeed, node balancing and face balancing don't change x'; so if no excitation occurred in step $t+1$, then $x'(t+1) - x'(t) = 0$, and new flow was created on edge e, then it depends only on e. Hence the vectors $y'(t+1) - y'(t)$ are also mutually independent.

Since $y'(t+1) - y'(t) = 0$ with probability $1 - p$, it makes sense to aggregate $N = 1/p$ of these terms to one. So we consider the vectors $z(t) = y(Nt) - y(N(t-1))$ and $z'(t) = y'(Nt) - y'(N(t-1))$. Then the vectors $z'(t)$ are mutually independent samples from some distribution on L. (The errors $z''(t) = z(t) - z'(t)$ may be dependent.)

A further difficulty is that if an excitation step occurs close to the end of an aggregated interval $[N(t-1)+1, Nt]$, then the error $z''(t)$ can be larger than the main term $z'(t)$. This happens with small but not negligible probability. We handle this by randomly selecting just a fraction of these intervals, so that the probability of any of these bad large errors occurring is small.

To formalize, we propose the following algorithm to recover the dimension of $\mathcal{C}$.

Genus estimate. We assume that we are given upper bounds $\overline{n} \geq n+m+f$ and $\overline{g} \geq g$. Let m_0 be the number of edges incident with U. Set

$$T' = 6(\overline{n} + \overline{g}), \qquad \varepsilon = \overline{n}^{-m_0 \overline{n}}, \qquad T = 4T'^2.$$

Construct a sequence of integers $t_1, t_2, \ldots, \in [0, T-1]$ as follows. If we have $t_1, t_2, \ldots, t_k$, then compute the linear hull $\mathcal{L}(k)$ of $z(t_1), z(t_2), \ldots, z(t_k)$. Let $H(k)$ be the set of integers $t \in [0, T-1]$ for which the vector $z(t)$ is farther from $\mathcal{L}(k)$ than ε. If $|H(k)| < T'$, then return $k/2$ as your guess for g. Else, choose a number $t_{k+1} \in H(k)$ randomly and uniformly. If k becomes larger that $2\overline{g}$, declare the procedure a failure and stop.

The main result of this paper is the following.

Theorem 4 *The Genus Estimate Algorithm returns the correct genus with probability at least* $2/3$.

If you find that a success probability of $2/3$ is not reassuring enough, independent repetition of the observation can boost this arbitrarily close to 1.

2 Properties of smooth circulations

2.1 Harmonic functions

It is an easy well-known fact that if $G = (V, E)$ is a connected graph (the orientation is not needed right now), then for any two nodes $a, b \in V$ there is a vector $\pi \in \mathbb{R}^V$ such that for every node v,

$$\sum_{u:\ uv \in E} \pi_u - d_v \pi_v = \begin{cases} 1, & \text{if } v = b, \\ -1, & \text{if } v = a, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

This last expression is equivalent to saying that $f_e = \pi_{h(e)} - \pi_{t(e)}$ is a flow. We denote this vector π by $\pi_{a,b}$ if we want to express that it depends on a and b; if there is an edge e with $h(e) = b$ and $t(e) = a$, then we also denote $\pi_{a,b}$ by π_e.

The vector π is not unique; we can add the same scalar to each entry. For our purposes, it will be convenient to choose it so that

$$\sum_u \pi_u = 0. \tag{2}$$

There are many interpretations of these functions; for example, let the graph represent an electrical network, with the edges having unit resistance. Send a unit electric current from b to a. Then the potential of u is π_u. The function $\pi_{a,b}$ is often called a *harmonic function with poles a and b*.

In terms of the Laplacian L of the graph, the equations (1) can be written as

$$L\pi = \chi_b - \chi_a.$$

The matrix L is not quite invertible, but it has a one-dimensional nullspace spanned by the vector $\mathbf{1} = (1, \dots, 1)^{\mathsf{T}}$, and so it determines π up to adding the same scalar to every entry. We assumed in (2) that $\mathbf{1}^{\mathsf{T}}\pi = 0$. If $J \in \mathbb{R}^{V \times V}$ denotes the all-1 matrix, then

$$(L + J)\pi = L\pi = \chi_b - \chi_a,$$

and so we can express π as

$$\pi = (L + J)^{-1}(\chi_b - \chi_a). \tag{3}$$

We can use harmonic functions to give a more explicit description of smooth circulations in a special case. For any edge e of G, let η_e be the projection of χ_e onto $\mathcal{C}$.

Lemma 5 *Let $a, b \in E$ be two edges of G. Then*

$$(\eta_a)_b = \begin{cases} (\pi_b)_{h(a)} - (\pi_b)_{t(a)} + (\pi_b^*)_{r(a)} - (\pi_b^*)_{l(a)} + 1, & \text{if } a = b, \\ (\pi_b)_{h(a)} - (\pi_b)_{t(a)} + (\pi_b^*)_{r(a)} - (\pi_b^*)_{l(a)}, & \text{otherwise.} \end{cases}$$

Proof. Let x_1, x_2 and x_3 be the projections of χ_b on the linear subspaces $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, respectively. The vector x_1 can be expressed as a linear combination of the vectors δv ($v \in V$), which means that there is a vector $y \in \mathbb{R}^V$ so that $x_1 = My$. Similarly, we can write $x_2 = Nz$. Together with $x = x_3$, these vectors satisfy the following system of linear equations:

$$\begin{cases} x + My + Nz = \chi_b \\ M^{\mathsf{T}} x = 0 \\ N^{\mathsf{T}} x = 0 \end{cases} \tag{4}$$

Multiplying the first m equations by the matrix M^{T}, and using the second equation and the fact that $M^{\mathsf{T}} N = 0$, we get

$$M^{\mathsf{T}} M y = M^{\mathsf{T}} \chi_b, \tag{5}$$

and similarly,

$$N^{\mathsf{T}} N z = N^{\mathsf{T}} \chi_b. \tag{6}$$

Here $M^{\mathsf{T}} M$ is the Laplacian of G and $N^{\mathsf{T}} N$ is the Laplacian of G^*, and so (5) implies that $y = \pi_b + c\mathbf{1}$ for some scalar c. Similarly, $z = \pi_b^* + c^*\mathbf{1}$ for some scalar c'. Thus

$$x = \chi_b - M^{\mathsf{T}}(\pi_b + c\mathbf{1}) - N^{\mathsf{T}}(\pi_b^* + c^*\mathbf{1}) = \chi_b - M^{\mathsf{T}}\pi_b - N^{\mathsf{T}}\pi_b^*,$$

which is just the formula in the lemma, written in matrix form. □

2.2 Nondegeneracy properties of smooth circulations

We state and prove two key properties of smooth circulations: one, that the projection of a basis vector to the space of smooth circulations is non-zero, and two, that smooth circulations are spread out essentially over the whole graph in the sense that every connected piece of the graph where a non-zero smooth circulation vanishes can be isolated from the rest by a small number of points.

Theorem 6 *If $g > 0$, then for every edge e, $\eta_e \neq 0$.*

Proof. Suppose that $\eta_e = 0$. Then by Lemma 5, there are vectors $\pi = \pi(e) \in \mathbb{R}^V$ and $\pi^* = \pi^*(e) \in \mathbb{R}^F$ such that

$$\pi_{h(a)} - \pi_{t(a)} = \pi^*_{r(a)} - \pi^*_{l(a)} \tag{7}$$

for every edge $a \neq e$, but

$$\pi_{h(e)} - \pi_{t(e)} = 1 + \pi^*_{r(e)} - \pi^*_{l(a)}. \tag{8}$$

For convenience, we orient every edge so that $\pi^*_{r(a)} \geq \pi^*_{l(a)}$.

Let $\alpha_1 < \alpha_2 < \ldots < \alpha_k$ be the different values of π^*, and let $\mathcal{F}_i = \{F \in \mathcal{F} : \pi^*_F = \alpha_i$. Since for every face F other than $r(e)$ and $l(e)$ the value of π^*_F is the average of its values on the neighbors, it follows that $\mathcal{F}_1 = \{l(e)\}$ and $\mathcal{F}_k = \{\mathcal{F} \setminus r(e)\}$.

Consider the union S_i of all faces in $\mathcal{F}_1 \cup \ldots \cup \mathcal{F}_i$, $1 \leq i < k$. We claim that its boundary is a single cycle containing the edge e. It is trivial that $e \subseteq \partial S_i$. Suppose that the boundary of S_i contained a cycle C that did not go through e, which is a directed cycle with S_i being on its left hand side. But then $\pi^*_{r(a)} > \pi^*_{l(a)}$ for every edge of C. By (7), this means that $\pi_{h(a)} > \pi_{t(a)}$ for every edge a of C, which is clearly impossible.

Thus we know that ∂S_i consists of a single cycle through e; in other words, it consists of e and a path P_i connecting $t(e)$ and $h(e)$. As above, it follows that P_i is a directed path from $h(e)$ to $t(e)$, and π_v strictly increases as we traverse the path.

Consider the two paths P_{i-1} and P_i. These may have common nodes besides their endpoints, but from the fact that π strictly increases along both of them, it follows that their common nodes are in the same order on both paths. This implies that the two paths have the following structure: there are common (directed) subpaths $Q_0, Q_1, \ldots, Q_r$ (possibly consisting of just a single node), and cycles $C_1, C_2, \ldots, C_r$ so that C_j is the union of two directed paths R_j and R'_j, connecting the endpoint of Q_{i-1} to the starting point of Q_i, with $R_j \subseteq P_{i-1}$ and $R'_j \subseteq P_i$.

The union of all faces $F \in \mathcal{F}_i$ is a (not necessarily connected) surface S'_i whose boundary is the union of the cycles C_j.

Claim. *S'_i consists of r discs $D_1, \ldots D_r$, where the boundary of D_j is the cycle C_j.*

Let G'_i be the subgraph of G contained in S'_i (including the nodes on the boundary, but not the edges). Let H be a connected component of G'. Since every edge of G' has $\pi_{h(a)} = \pi_{t(a)}$ by (7), all nodes of H have the same π value. Since π is strictly increasing along P_{i-1} as well as along P_i, it follows that H can be attached to at most one node on each of these paths. By the 3-connectivity of G, it follows that H is a single edge connecting some $u \in V(P_{i-1})$ and $v \in V(P_i)$. From $\pi_u = \pi_v$ it also follows that H connects a node on an R_j to a node on the corresponding R'_j, and the endpoints of these edges are in the same order on both paths, which gives an ordering of the edges of G'.

Now every face in $\mathcal{F}_i$ is a disc attached along a cycle in $P_{i-1} \cup P_i \cup G'$, so that the edges in G' have two faces attached, while the edges of $P_{i-1} \cup P_i$ have only one. The only way to this is to attach each face in $\mathcal{F}_i$ to

— two consecutive edges of G' and to the to subpaths of G between their endpoints, or

— to the first or last edge of G' connecting two points on a cycle C_j and to the corresponding arc of C_j.

Hence the Claim follows immediately.

Now we see that S_i is obtained from S_{i-1} by attaching the disjoint disks $D_1, \ldots, D_r$ along single arcs $R_1, \ldots, R_r$. Since S_1 is a disc, it follows by induction that S_i is a disk, and in particular, S_{k-1} is a disc. Since S is obtained from S_{k-1} by gluing in the last face (a disc) along the cycle ∂S_{k-1}, it follows that S is a sphere, and so $g = 0$. □

Theorem 7 *Let G be a graph embedded in an orientable surface S of genus g so that all faces are discs. Let h be a non-zero smooth circulation on G and let G' be the subgraph of G on which h does not vanish. Suppose that h vanishes on all edges incident with a connected subgraph U of G. Then U can be separated from G' by at most $16g$ points.*

The assumption that the connectivity between U and the rest of the graph must be linear in g is "essentially" sharp in the following sense. Suppose that G has a cutset of $2g-1$ or fewer edges. Since the dimension of the space of smooth circulations is $2g$, there will be a non-zero smooth circulation h vanishing on these edges. If this circulation is non-zero on (say) the left hand side of this cut, then we can replace all

Before the proof of this theorem we need a simple lemma about maps.

Lemma 8 *Let G be any digraph embedded on an orientable surface S of genus g. Assume that G has no sources and sinks. For every face F, let a_F denote the number of nodes v on the boundary of F for which the two edges on the boundary of F incident with v are directed to v. Then*

$$\sum_F (a_F - 1) \le 2g - 2.$$

Proof. Let n, m and f denote the number of nodes, edges and faces. Clearly a_F is also the number the nodes on the boundary of F with both edges oriented out, and so $\sum_F 2a_F$ counts the number of "corners" with both edges oriented in or both edges oriented out. Since there are no sources or sinks, at every node there are at least 2 corners with one edge in and one out, and hence

$$\sum_F 2a_F \le \sum_v (d_v - 2) = 2m - 2n = 2f + 4g - 4.$$

Rearranging and dividing by 2, we get the inequality of the lemma. □

Corollary 9 *If there is no face whose boundary is a directed cycle, then $a_F \le 2g - 1$ for every face.*

Now we are ready to prove Theorem 7.

Proof. By re-orienting edges of G' we may assume that $h > 0$ on the edges G'. Let us shrink every connected component of $G \setminus V(G')$ to a single point, to get a digraph G'' embedded in S. Note that every face of G'' is a disc and it contains at least one edge of G', and so it cannot be bounded by a directed cycle.

Let w be the node to which U is contracted, let F' be the face of G' containing w. We think of F as a surface with a boundary ∂F consisting of k disjoint Jordan curves $C_1, \dots, C_k$, glued to some (not necessarily disjoint) cycles of the graph G'.

Let R be a face of G'' contained in F', bounded by a cycle ∂R. Then $\partial F \cap \partial R$ consists of one or more arcs; let b_R denote the number of these arcs.

Let r_i denote the number of faces R with $b_R = i$, and let r'_i be the number of faces among these which touch w.

First we estimate r_1. Along the common arc of ∂F and ∂R, the direction of the edges must change at least once: else, h would add up to a non-zero value along the face boundary ∂R, contradicting the assumption that h is smooth. Hence by Lemma 9, $r_1 \le 2a_{F'} \le 4g - 2$.

Consider a face R with $b_R > 1$ touching w. Select a node v_R in the interior of R, and connect it inside R by disjoint arcs to w and to one point on each of the arcs of $\partial F \cap \partial R$. Call these arcs *red*.

Cut away all of the surface outside F', and contract each C_i to a single point. Also contract W to a single point. This way we get a surface S^* with genus g^*. The red arcs form a graph G^* embedded in S^*. It is clear that G^* is bipartite.

Next, we estimate the number f_2^* digonal faces of G^*. Indeed, a digonal face must be formed by two red arcs originally connecting a node v_R to two points a and a' on different common arcs A and A' of ∂R and a C_i, and by a disc R' bounded by these two red arcs and an arc B of C_i connecting a and a'. Since A and A' are distinct arcs, there must be a face of G inside R' attached to at least one edge of C_i. In fact, there must be such a face R_0 that is attached to C_i along a single arc. Clearly, R_0 is disjoint from W and so it is counted in $r_1 - r_1'$. It is also clear that different digonal faces of G^* correspond to different faces counted in $r_1 - r_1'$. This proves that $f_2^* \leq r_1 - r_1'$.

Now use Euler's formula. The number of nodes of G^* is $1+k+k'$, where the 1 accounts for w, k is the number of components C_i, and $k" = \sum_{i\geq 2} r_i'$ is the number of faces of G meeting W. The number of edges is $m^* = \sum_{i\geq 2}(i+1)r_i'$. So the number of faces is

$$f^* = \sum_{i\geq 2}(i+1)r_i' - \left(1 + k + \sum_{i\geq 2} r_i'\right) + 2g^* - 2.$$

There are f_2^* digons and $f^* - f_2^*$ faces that are all at least 4-gons, and hence

$$2m^* \geq 2f_2^* + 4(f* - f_2^*) = 4f^* - 2f_2^* \geq 4f^* - 2(r_1 - r_1') \geq 4f^* - 4g - 2 + 2r_1',$$

and so

$$\sum_{i\geq 2}(i+1)r_i' = m^* \geq 2f^* - 4g - 2 + r_1' = r_1' + \sum_{i\geq 2} 2ir_i' - 2k - 2 - 4g^* + 4 - 4g + 2$$

whence

$$r_1' + \sum_{i\geq 2}(i-1)r_i' \leq 2k + 4g^* + 4g - 4.$$

Let N denote the number of neighbors of w on ∂F. Each neighbor of w is incident with at least two faces of G'' containing w and an arc of ∂F, and a face containing i arcs of ∂F is counted at most $2i$ times this way; hence

$$2N \leq \sum_{i\geq 1} 2ir_i' \leq 4\left(r_1' + \sum_{i\geq 2}(i-1)r_i'\right) \leq 8k + 16g + 16g^* - 16.$$

To conclude the proof, it suffices to show that

$$2g * + k \leq 2g \tag{9}$$

unless $g = g^* = 0$ and $k = 1$. To this end, notice that no cycle C_i bounds a disk not containing w. Indeed, the restriction of h to the connected component of G' contained in this disk would be a smooth circulation on a planar map, contradicting Lemma 3.

If $k = 1$ and C_1 bounds a disc containing w, then $g^* = 0$ and (9) is trivial. In every other case, none of the cycles C_i is null-homologous.

Let us cut the surface S along each C_i one by one, and in each case, glue discs on both copies of C_i obtained by cutting. If this operation separates the surface, then we keep the part containing w only. This way we obtain the surface S^*.

Every cut reduces the genus of the surface, except when it separates the surface and the part thrown away is a sphere. Since each such part must have been attached along at least two cycles C_i, at least half of the cuts reduce the genus, showing that $g^* \le g - k/2$. This proves (9) and completes the proof of Theorem 7. □

3 Proof of Theorem 4

3.1 A lemma about convergence

We prove a simple lemma about the convergence of a simple randomized iterative projection process.

Let $a_1, \dots, a_k \in \mathbb{R}^n$, and let $A = (1/k)\sum_{i=1}^k a_i a_i^\mathsf{T}$. Let λ be the smallest positive eigenvalue of A. Let L be the linear subspace generated by $a_1, \dots, a_k$, and $L^\perp$, its orthogonal complement.

For $x \in \mathbb{R}^n$, define a Markov chain $X^0, X^1, \dots \in \mathbb{R}^n$ as follows: start with $X^0 = x$. Given X^t, choose a vector a_i (uniformly and randomly), and let

$$X^{t+1} = X^t - \frac{a_i^\mathsf{T} X^t}{a_i^\mathsf{T} a_i} a_i.$$

Let y and z be the orthogonal projections of x onto $L^\perp$ and L, respectively. Then $X^t \to y$, and in fact the following lemma describes the rate of convergence:

Lemma 10 *For every $x \in \mathbb{R}^n$,*

$$\mathsf{E}\left(|X^t - y|^2\right) \le (1 - \lambda)^t |z|^2.$$

Proof. First we consider the case $t = 1$:

$$|X^1 - y|^2 = \left| z - \frac{a_i^\mathsf{T} z}{a_i^\mathsf{T} a_i} a_i \right|^2 = |z|^2 - \frac{(a_i^\mathsf{T} z)^2}{|a_i|^2},$$

and hence

$$\mathsf{E}\left(|X^t - y|^2\right) = |z|^2 - z^\mathsf{T} \mathsf{E}\left(\frac{1}{a_i^\mathsf{T} a_i} |a_i|^2\right) z$$

$$= |z|^2 - z^\mathsf{T} \left(\frac{1}{k} \sum_{i=1}^{k} \frac{1}{|a_i|^2} a_i a_i^\mathsf{T}\right) z = |z|^2 - z^\mathsf{T} A z.$$

Since $z \in L$ is in the range of A, we have $z^\mathsf{T} A z \geq \lambda |z|^2$, which proves the assertion.

The general case follows by induction. □

3.2 Setup for the proof

Let $a_i \in \mathbb{R}^E$ be defined for every node i by

$$(a_i)_e = \begin{cases} 1, & \text{if } i \text{ is the head } e, \\ -1, & \text{if } i \text{ is the tail of } e, \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, let $b_F \in \mathbb{R}^E$ be defined for every face F by

$$(b_F)_e = \begin{cases} 1, & \text{if } e \text{ is an edge of } \partial F \text{ oriented clockwise}, \\ -1, & \text{if } e \text{ is an edge of } \partial F \text{ oriented counterclockwise}, \\ 0, & \text{otherwise.} \end{cases}$$

We consider the matrices

$$A = \frac{1}{n} \sum_{i \in V} \frac{1}{|a_i|^2} a_i a_i^\mathsf{T},$$

and

$$B = \frac{1}{f} \sum_{F \in \mathcal{F}} \frac{1}{|b_F|^2} b_F b_F^\mathsf{T}.$$

Let λ_1 and λ_2 be the smallest positive eigenvalue of A and B, respectively, and let $\mu = \min\{\lambda_1, \lambda_2\}$.

Let $T_t : \mathbb{R}^E \to \mathbb{R}^E$ denote the (random) linear mapping that (a) and (b) generate in step t. Note that subspace $\mathcal{C}$ is invariant under T_t. Let $W(t,s) = T_s T_{s-1} \dots T_{t+1}$,

$$u(t) = \begin{cases} e_j, & \text{if in step } t \text{ edge } j \text{ was excited,} \\ 0, & \text{otherwise.} \end{cases}$$

and $u(t,s) = W(t,s)u(t)$. Hence

$$x(s) = \sum_{t=0}^{s} u(t,s).$$

Let $u_1(t,s)$, $u_2(t,s)$ and $u_3(t,s)$ denote the orthogonal projections of $u(t,s)$ to $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, respectively, and let $u_i(t) = u_i(t,t)$. This notation is somewhat redundant, since $\mathcal{C}$ is invariant under T_t, and hence $u_3(t,s) = u_3(t)$ for every $s \geq t$. The "error part" is $w(t,s) = u_1(t,s) + u_2(t,s)$. Thus we get the "smooth part"

$$x'(s) = \sum_{t=0}^{s} u_3(t),$$

and the "error part"

$$x''(s) = \sum_{t=0}^{s} w(t,s). \tag{10}$$

We need to bound x' from below and x'' from above.

3.3 Bounding smooth circulations

The vector $u_3(t)$ is a smooth circulation. Suppose that it is not the 0 circulation. Then by Theorem 7, we know that its restriction $v(t)$ to E_0 is not the 0 circulation. We need a lower bound on $|v(t)|$. Recall that $f = |\mathcal{F}|$; then we have

Lemma 11

$$|v(t)| \geq n^{-n} f^{-f}.$$

Proof. x is a rational vector whose denominator is a divisor of $\det(M^{\mathsf{T}}M + J)\det(N^{\mathsf{T}}N + J)$. By Hadamard's inequality, the denominator of x is at most $n^n f^f$.

Now $v(t)$ is a restriction of x which is nonzero, and so at least one coordinate of $v(t)$ is a non-zero rational number with denominator at most $n^n f^f$. This proves (11). □

3.4 Bounding the error

We prove a bound on the "error term". More exactly, we fix an integer $a > 0$, and split the error into "old errors" and "new errors":

$$x''(s) = \sum_{t=0}^{s-a} w(t,s) + \sum_{t=s-a+1}^{s} w(t,s) = X_1(s,a) + X_2(s,a).$$

First we estimate the expectation of $|X_1(s)|^2$. We claim that

$$\mathsf{E}(|X_1(s,a)|^2) \le 5\frac{p}{\mu}(1-\mu)^a. \tag{11}$$

Let us fix the "excitations" $u(0), u(1), \dots$, and let E_u denote expectation conditional on these. Lemma 10 implies that

$$\mathsf{E}_u(|u_1(t,s)|^2) \le (1-\lambda_1)^{s-t}|u_1(t)|^2,$$

and

$$\mathsf{E}_u(|u_2(t,s)|^2) \le (1-\lambda_2)^{s-t}|u_2(t)|^2,$$

and so

$$\begin{aligned} \mathsf{E}_u(|w(t,s)|^2) &= \mathsf{E}_u(|u_1(t,s)|^2 + |u_2(t,s)|^2) \qquad (12)\\ &\le (1-\mu)^{s-t}(|u_1(t)|^2 + |u_2(t)|^2)\\ &\le (1-\mu)^{s-t}|u(t)|^2. \qquad (13) \end{aligned}$$

From the definition of Σ_1 we have

$$\mathsf{E}_u(|\Sigma_1(s)|^2) = \mathsf{E}_u\left(\left|\sum_{t=0}^{s-a} w(t,s)\right|^2\right)$$

$$= \mathsf{E}_u\left(\sum_{t=0}^{s-a}\sum_{t'=0}^{s-a} w(t,s)^{\mathsf{T}} w(t',s)\right) = \sum_{t=0}^{s-a}\sum_{t'=0}^{s-a} \mathsf{E}_u(w(t,s)^{\mathsf{T}} w(t',s))$$

$$\le \sum_{t=0}^{s-a}\sum_{t'=0}^{s-a} \mathsf{E}_u(|w(t,s)|^2)^{1/2}\mathsf{E}_u(|w(t',s)|^2)^{1/2} = \left(\sum_{t=0}^{s-a} \mathsf{E}_u(|w(t,s)|^2)^{1/2}\right)^2.$$

Using (12), this gives

$$\mathsf{E}_u(|\Sigma_1(s)|^2) \le \left(\sum_{t=0}^{s-a}(1-\mu)^{(s-t)/2}|u(t)|\right)^2. \tag{14}$$

Now we take expectation over the sequence $u(t)$. Since the $|u(t)|$ are independent 0-1 valued variables with mean p, the expectation of the right hand side is easy to estimate:

$$\mathsf{E}\left(\left(\sum_{t=0}^{s-a}(1-\mu)^{(s-t)/2}|u(t)|\right)^2\right)$$

$$=\sum_{t=0}^{s-a}\sum_{t'=0}^{s-a}(1-\mu)^{(s-t)/2}(1-\mu)^{(s-t')/2}\mathsf{E}(|u(t)||u(t')|).$$

Here

$$\mathsf{E}(|u(t)||u(t')|)=\begin{cases}p, & \text{if } t=t',\\ p^2, & \text{otherwise.}\end{cases}$$

Thus we can write the sum above as

$$p^2\sum_{t=0}^{s-a}\sum_{t'=0}^{s-a}(1-\mu)^{(s-t)/2}(1-\mu)^{(s-t')/2}+(p-p^2)\sum_{t=0}^{s-a}(1-\mu)^{s-t},$$

$$=p^2(1-\mu)^a\left(\frac{1-(1-\mu)^{(s-a)/2}}{1-(1-\mu)^{1/2}}\right)^2+(p-p^2)(1-\mu)^a\frac{1-(1-\mu)^a}{1-(1-\mu)}$$

$$\le (1-\mu)^a\frac{4p^2}{\mu^2}+\frac{p-p^2}{\mu}<5\frac{p}{\mu}(1-\mu)^a.$$

This proves (11).

Second, we consider X_2. If no excitation event occurs between times $s-a+1$ and s, then $\Sigma_2(s)=0$. Hence

$$\Pr(|\Sigma_2(s)|>0)\le 1-(1-p)^a<ap. \tag{15}$$

From (11) and (15) we get that

$$\Pr(|x''(s)|>\delta)\le\Pr(|X_1(s)|\ge\delta)+\Pr(|X_2(s)|>0)<\frac{5p}{\mu\delta^2}(1-\mu)^a+ap.$$

The choice $a=\frac{2}{\mu}\ln\frac{1}{\delta}$ gives the best bound (up to a constant):

$$\Pr(|x''(s)|>\delta)<\frac{10p}{\mu}\ln\frac{1}{\delta}. \tag{16}$$

3.5 Completing the proof

To simplify notation, put $z_i = z(t_i)$, $z'_i = z'(t_i)$ and $z''_i = z''(t_i)$. It follows from the choice of the integers t_i that the vectors $z_1, z_2, \dots z_k$ are linearly independent, and so $\dim(\mathcal{L}(i)) = i$. Furthermore, by the selection of these vectors, the Gram-Schmidt orthogonalization $z_1^* = z_1, z_2^*, \dots, z_k^*$ consists of vectors of length at least ε.

Let $\delta = (\varepsilon/4)(1 + 1/\varepsilon)^{-2g_0}$. Since $|z''(t)| \le |x''(t)|$, the probability that $|z''(t)| > \delta$ is less than $\frac{10p}{\mu} \ln \frac{1}{\delta}$ by (16), so the probability that any of the t_i have $|z''(t_i)| > \delta$ is less than

$$g_0 \frac{10p}{\mu} \ln \frac{1}{\delta} < \frac{1}{9}.$$

So with probability at least .99, we have $|z''_i| \le \delta$ for every i. Let us assume that this occurs.

Claim 1. Suppose that for some real numbers $\alpha_1, \dots, \alpha_k$, we have

$$\left| \sum_{i=1}^{k} \alpha_i z_i \right| \le 1. \tag{17}$$

Then

$$|\alpha_i| \le \frac{1}{\varepsilon} \left(1 + \frac{1}{\varepsilon}\right)^{k-i}.$$

Indeed, (17) implies that for any $1 \le j \le k$,

$$\left| \sum_{i=1}^{k} \alpha_i z_i^{\mathsf{T}} z_j^* \right| \le |z_j^*|.$$

Since $z_i^{\mathsf{T}} z_j^* = 0$ if $i < j$ and $z_j^{\mathsf{T}} z_j^* = |z_j^*|^2$, it follows that

$$|\alpha_j||z_j^*|^2 \le |z_j^*| - \sum_{i=j}^{k} |\alpha_i z_i^{\mathsf{T}} z_j^*| \le |z_j^*| + \sum_{i=j}^{k} |\alpha_i||z_j^*|.$$

Dividing by $|z_j^*|^2$ and using that $|z_i| \le 1$, we get that

$$|\alpha_j| \le \frac{1}{|z_j^*|} \sum_{i=j}^{k} |\alpha_i| \le \frac{1}{\varepsilon} \sum_{i=j}^{k} |\alpha_i|.$$

Hence the claim follows by induction on $k - i$.

Claim 1. The vectors $z'_1, \dots, z'_{2g}$ are linearly independent.

Indeed, assume that there is a linear relation $\sum_i \alpha_i z'_i = 0$ where not all the α_i are 0. We can write this as $\sum_i \alpha_i z_i = \sum_i \alpha_i z''_i$. Since the z_i are linearly independent, the left hand side is non-zero, and so we may assume it has norm 1. But then Claim 1 implies that $|\alpha_i| \le (1/\varepsilon)(1+1/\varepsilon)^{k-i}$, and so using that $|z_i| \le \delta$,

$$\left|\sum_i \alpha_i z''_i\right| \le \delta \sum_i |\alpha_i| \le \delta \left(1 + \frac{1}{\varepsilon}\right)^k < 1,$$

a contradiction.

Claim 3. $d(z(t), \mathcal{L}(k)) < |z''(t)| + \frac{\varepsilon}{2}$.

Indeed, by Claim 2 we can write $z'(t) = \sum_{i=1}^k \alpha_i z'_i$ with some real numbers α_i. Then

$$\sum_{i=1}^k \alpha_i z_i = z'(t) + \sum_{i=1}^k \alpha_i z''_i.$$

Let R be the norm of the vector on the two sides. Then by Claim 1, we have $|\alpha_i| \le R(1/\varepsilon)(1+1/\varepsilon)^{k-i}$. Using this, we get

$$R = \left|z'(t) + \sum_{i=1}^k \alpha_i z''_i\right| \le 1 + R\delta\left(1 + \frac{1}{\varepsilon}\right)^k \le 1 + \frac{1}{2}R,$$

whence $R \le 2$, and so $|\alpha_i| \le (2/\varepsilon)(1+1/\varepsilon)^{k-i}$.

Let $w = \sum_{i=1}^k \alpha_i z_i \in \mathcal{L}(k)$. Then

$$|w - z'(t)| = \left|\sum_{i=1}^k \alpha_i z''_i\right| \le \delta \sum_{i=1}^k |\alpha_i| \le 2\delta\left(1 + \frac{1}{\varepsilon}\right)^k < \varepsilon/2,$$

and so

$$d(z(t), \mathcal{L}(k)) \le |z(t) - w| \le |z(t) - z'(t)| - |z'(t) - w| < |z''(t)| + \frac{\varepsilon}{2},$$

as claimed.

Now we are ready to complete the proof. In guessing that $k/2$ is the genus of the surface, we can err in two directions: it may be that $g < k/2$ or $g > k/2$. We estimate the probability of these errors separately.

If $g < k/2$, then we moved on when we had $k = 2g$, which means that at that stage we had too many vectors $z(t)$ farther from $\mathcal{L}(k)$ than ε. By

Claim 3, every such t must satisfy $|z''(t)| > \varepsilon/2 > \delta$. By (16), the probability that a given t has this property is less than $\frac{10p}{\mu} \ln \frac{1}{\delta}$, and so using Chernoff's inequality, the probability that the number of indices t with this property is at least T' is less than 1/9.

If $g > k/2$, then we stopped too early: there was a subspace L' with $\dim(L') < \dim(L)$ so that almost all the $z(t)$ were closer to L' than n^{-10n}. The mapping $\mathbb{R}^E \to \mathcal{L}$ obtained by projecting $\mathbb{R}^E$ to $\mathcal{C}$ and then restricting it to E_0 is surjective; this implies that there are $2g$ edges $e_1, \dots, e_{2g} \in E$ so that the vectors $\mu(e_1), \dots, \mu(e_{2g})$ form a basis in $\mathcal{L}$. These vectors are rational with denominators at most $n^n f^f$, so the determinant of this basis is at least $n^{-m_0 n} f^{-m_0 f} \geq \varepsilon$. It follows that at least one of these vectors, say $\mu(e_1)$, is at distance at least ε from $\mathcal{L}$.

Thus if in any of the time intervals $[(t-1)N+1, tN]$, the edge e_1 was excited, no other edge was excited, and the excitation of e_i occurred in the first $N-a$ steps in this time interval, then $z(t)$ must be farther from $\mathcal{L}'$ than ε. The probability that this happens for a given t is $(N-a)p(1-p)^{N-1}/m$; so (using the Chernoff bound again) the probability that this happens for fewer than T' choices of t is less that 1/9.

To sum up, the total probability of "bad" cases is less than 1/9 (when $|z_i| > \delta$ for some i) plus 1/9 (when $g < k/2$) plus 1/9 (when $g > k/2$). This proves the theorem.

4 Concluding remarks

1. We can make some cosmetic changes to the setup as given above. One objection may be that the noisy circulator, as defined, is not truly local, since (say) in operation (a) we have to select a node uniformly from all nodes. The standard way of fixing this is to attach an "alarm clock" to each node, edge, and face, which wakes them up at random times according to a Poisson process (the edge-clock is much slower than the other two).

Another objection is that the noisy circulator as constructed above is not stationary: the total mass grows to infinity. An easy fix is to give a second, even slower clock to each edge: when this rings, they reset their value to 0. Another possible fix comes from the observation that the excitations don't necessarily have to be constants, the proof works just as well for random and symmetric excitations. So modifying the excitation step to reset the value of any edge to 1 with very small probability (rather than to add 1) provides a stationary version (but the analysis becomes more complicated).

Further variants, improvements and generalizations of the above system are of interest:

- Can one recover the genus even if the rate at which excitations take place is faster (ideally, independent of the number of nodes)?
- Suppose that we only allow two values (or any other given discrete set of values) on any edge. Can we still recover the genus?
- Can one extend the technique to recover a non-orientable surface by observing a random process on an embedded graph locally?

For background regarding graphs on surfaces and Riemann surfaces see for instance [11] [4], for background regarding algorithmic applications of random processes see [7].

2. The notion of global information from local observation can inspire many questions in different directions. We briefly present some related examples.

Example 1 Let G be a finite connected graph. Start a simple random walk on G. Fix a vertex v in G. You are given the sequence of times for which the simple random walk visits v, what information can be learned about G? From the infinite sequence one can reconstruct the on diagonal heat kernel and thus the spectrum of the transition matrix of the random walk (so in the case of regular graphs, the spectrum of the graph).

Example 2 [Obtaining the size of the road system by measuring the volume of local traffic]. Let G be a finite connected irreducible regular graph. From each vertex start an independent simple random walk. Fix a vertex $v \in G$. For each time t you can observe the number of walkers occupying v. How much time is needed in order to, almost surely, know the size of G? Assume you are given an *a priori* bound N on $n = |V(G)|$. Then one can get a (poor) polynomial upper bound on the time needed along the following lines. The mixing time of G is smaller than n^3, so if we observe the load on v only at times of the form kN^3, we get almost independent samples from a distribution which is exponentially close to Binomial$(n, 1/n)$, the stationary distribution for the number of walkers at v. For Binomial$(n, 1/n)$, about n^4 samples are needed to recover n. So after about N^7 steps n can be recovered. (Probably N^4 is the time needed to recover n.)

There are several random processes on graphs which have been considered before, and for which the question "what global information can be deduced from local observation" is meaningful. We mention two examples.

Example 3 Consider an n-vertex connected graph. There are $k \leq n$ particles labeled $1, 2, ..., k$. In a configuration, there is one particle at each vertex. The interchange process discussed briefly in [1], is a continuous-time Markov chain on configurations. For each edge (i, j), at rate 1 the particles at vertex i and vertex j are interchanged. If only one of the two vertices is occupied, then it jumps to the other vertex. Assume you observe which labeled particle occupies a fixed predetermined vertex at any time. For $k = 1$, this is just Example 1. If $k > 1$, can one recover, using these observations, further information about G not contained in the spectrum? If $k = n$, is it possible to reconstruct G?

Example 4 Another natural candidate for local observation is the heat-bath chain (Glauber dynamics) on k-colorings of a graph: at any step, we have a (legal) k-coloring; we select a random node v and a random color α, and we re-color v with color α if this gives a legal k-coloring. Can we derive estimates on the chromatic number, or maximum degree, by observing a bounded piece of the graph?

Acknowledgement. We are grateful to Dimitris Achlioptas, Mike Freedman, Eran Makover, Oded Schramm and Kevin Walker for stimulating discussions on this topic.

References

[1] D. Aldous, My favorite 6 open problems in mathematical probability. Preprint, (2001).

[2] B. Bollobás, Volume estimates and rapid mixing. *Flavors of geometry*, 151–182, Math. Sci. Res. Inst. Publ., **31**, Cambridge Univ. Press, Cambridge, 1997.

[3] R.L. Brooks, C.A.B. Smith, A.H. Stone, W.T. Tutte, The dissection of rectangles into squares. Duke Math. J. 7, (1940). 312–340. (Reviewer: P. Scherk) 48.0X

[4] P. Buser, Geometry and spectra of compact Riemann surfaces. Progress in Mathematics, 106. Birkheuser Boston, Inc., Boston, MA, xiv+454 pp, 1992.

[5] R.J. Duffin, Basic properties of discrete analytic functions, *Duke Math. J.* **23** (1956), 335–363.

[6] J. Ferrand, Fonctions préharmoniques et fonctions préholomorphes. (French) *Bull. Sci. Math.* **68**, (1944). 152–180.

[7] O. Häggström, Finite Markov chains and algorithmic applications. Preprint, (2002).

[8] H. Kesten, Distinguishing and reconstructing sceneries from observations along random walk paths. Microsurveys in discrete probability (Princeton, NJ, 1997), 75–83, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI, 1998.

[9] L. Lovász and P. Winkler, Mixing times. Microsurveys in discrete probability (Princeton, NJ, 1997), 85–133, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI, 1998.

[10] C. Mercat, Discrete Riemann surfaces and the Ising model. *Comm. Math. Phys.* **218** (2001), no. 1, 177–216.

[11] B. Mohar and C. Thomassen, Graphs on surfaces. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, xii+291 pp, 2001.