Combinatorics, Probability and Computing (1997) 6, 27–38
Copyright © 1997 Cambridge University Press

On the Error-Correcting Capabilities of Cycle Codes of Graphs

LAURENT DECREUSEFOND and GILLES ZÉMOR

Ecole Nationale Supérieure des Télécommunications, 46 rue Barrault, 75 634 Paris Cedex 13, France

Received 2 September 1995; revised 15 December 1995

We are interested in a function $f(p)$ that represents the probability that a random subset of edges of a Δ-regular graph G contains half the edges of some cycle of G. $f(p)$ is also the probability that a codeword is corrupted beyond recognition when words of the cycle code of G are submitted to the binary symmetric channel. We derive a precise upper bound on the largest p for which $f(p)$ can vanish when the number of edges of G goes to infinity. To this end, we introduce the notion of fractional percolation on trees, and calculate the related critical probabilities.

1. Introduction

1.1 Cycle codes of graphs

Let $\mathbb{F} = \{0, 1\}$ denote the field on two elements. For any set S denote by 2^S the set of subsets of S. If $\mathbf{x}, \mathbf{y} \in 2^S$, $\mathbf{x}+\mathbf{y}$ will denote the symmetric difference of $\mathbf{x}$ and $\mathbf{y}$. 2^S is in a natural correspondence with $\mathbb{F}^s$, the vector space of binary s-tuples, where $s = \# S$. We shall frequently abuse notation by identifying subsets of S with their characteristic vectors in $\mathbb{F}^s$.

Let G be a finite graph. Denote by $V(G)$ and $E(G)$ the set of vertices and the set of edges of G, respectively, or simply V and E when no confusion can arise. Let $v = \# V$ and $n = \# E$ denote the cardinalities of V and E. Graphs will be undirected, without loops and multiple edges: i.e. an edge is an element of 2^V containing exactly two vertices. For any edge $e \in E$, define its *boundary* $\partial e \in 2^V$ as the union of its endpoints. ∂ is naturally extended to a mapping of 2^E to 2^V, where,

$$\partial : \mathbf{x} \mapsto \sum_{e \in \mathbf{x}} \partial e.$$

A *cycle* is a set of edges with zero boundary. We shall denote the subspace of $\mathbb{F}^n$ of cycles of G by $\mathscr{C}(G)$ and refer to it as the *cycle code* of G. If the graph G is connected, which we shall always suppose in what follows, $\mathscr{C}(G)$ has dimension $k = \dim \mathscr{C}(G) = n - v + 1$. We

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

shall say that a graph is Δ-regular if every vertex has exactly Δ neighbours. In this case, $k = \dim \mathscr{C}(G) = n(1-2/\Delta)+1$. For background on cycle codes of graphs, see [1]. The size of the smallest cycle in G is called the *girth* of G by graph-theorists, and is the *minimum distance* of $\mathscr{C}(G)$ for coding theorists: we shall denote it by $d(G)$, or simply d.

In the present work, we wish to judge the quality of cycle codes of graphs from the coding theory viewpoint. Even though one of the first things to be uncovered about cycle codes of graphs is that they have poor minimum distance properties for growing n (for instance, because, for a given rate k/n, $d(G)$ cannot grow faster than a logarithm of n), their simple structural appeal has attracted extensive study in the early days of coding theory (e.g. see [3]). Today, one of the remaining open questions about cycle codes of graphs that we wish to address is: 'For a given information rate $R = k/n$, what is the highest channel error probability that these codes can sustain, while achieving vanishing (with block length n) residual error probability after decoding?' Let us make this question more precise.

1.2 Error-probabilities and threshold behaviour

We are interested in the function $f_G(p)$ defined as

$$\begin{aligned}[0,1] &\to [0,1]\\ p &\mapsto f_G(p) = \sum_{\mathbf{x}\in W} p^{w(\mathbf{x})}(1-p)^{n-w(\mathbf{x})}\end{aligned}$$

where $w(\mathbf{x})$ denotes the weight of $\mathbf{x}$, and where

$$W = \{\mathbf{x}\in 2^E \mid \exists \mathbf{c}\in\mathscr{C}(G), \mathbf{c}\neq\mathbf{0}, w(\mathbf{x}\cap\mathbf{c}) \geqslant w(\mathbf{c})/2\},$$

i.e. W is the set of subsets of edges that contain at least half the edges of some non-zero cycle. In other words, $f_G(p)$ is the probability that a random set of edges $\mathbf{x}$ contains half the edges of some cycle, when $\mathbf{x}$ is obtained by choosing every edge independently with probability p.

From the coding point of view, we are submitting a codeword $\mathbf{c}'$ of $\mathscr{C}(G)$ to the binary symmetric channel with error-probability p. This means that each coordinate of $\mathbf{c}'$ is transformed into the complementary binary symbol independently with probability p. Equivalently, the received vector is $\mathbf{c}'+\mathbf{x}$, where $\mathbf{x}$ is some random error vector. The received vector is decoded by choosing the codeword closest to it for the Hamming distance. Whenever decoding produces a vector different from $\mathbf{c}'$ or a choice between two (or more) closest codewords, we shall say that a *decoding* (*or residual*) *error* occurs. One readily sees that the probability that a decoding error occurs is the probability that the error vector $\mathbf{x}$ contains half the coordinates of some parasite nonzero codeword $\mathbf{c}$, in other words, it is exactly $f_G(p)$ (actually, there is a small technical difference between $f_G(p)$, and what is classically named 'maximum likelihood decoding', but the difference will easily be seen to be insignificant asymptotically).

Let us define the parameter $\theta(G)$ associated with G as the real number such that $f_G(\theta) = 1/2$. We call θ the *threshold probability* of G. Such a definition is motivated by a theorem of Margulis, which proves that $f_G(p)$ tends to always display a threshold behaviour, i.e. jumps suddenly from almost zero to almost one; more precisely, we have the following interpretation of Margulis' theorem [5, 8].

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

Theorem 1. *For any* $\epsilon, \alpha > 0$, *there exists* $\delta > 0$ *such that whenever the girth* $d(G)$ *of* G *satisfies* $d > \delta$, *then*

$$\lambda(\{p, \epsilon < f_G(p) < 1-\epsilon\}) < \alpha,$$

where $\lambda(X)$ *denotes the Lebesgue measure of* X.

Let us define the threshold probability associated to a sequence of graphs $\mathscr{G} = (G_m)$ as $\theta(\mathscr{G}) = \limsup_{m\to\infty} \theta(G_m)$.

For any family of graphs $\mathscr{G} = (G_m)$ with fixed rate and girth $d(G_m)$ going to infinity with m (however slowly), the parameter $\theta(\mathscr{G})$ is a good measure of the error-correcting capabilities of the cycle codes of the graphs G_m for large m. It gives the maximum channel error probability p that members of G_m can sustain while maintaining vanishing residual error probability after decoding. θ is really, in an information-theoretic sense, a capacity notion, designed for codes rather than channels. Since the term 'capacity' is rather overemployed, we shall not use it, however.

In this paper, we strive to determine the largest possible threshold probability θ for cycle codes of Δ-regular graphs when m goes to infinity.

1.3 Trees

A basic remark is that as the girths of our graphs increase, locally they look like trees. We make this more formal by considering the following object. For a graph G and a vertex r, let us call the *covering tree of* G *rooted at* r, $\Gamma_r(G)$ (or simply, Γ_r), the infinite tree whose vertices are labelled by the paths of G starting at r. $\Gamma_r(G)$ has root r and the successors of a vertex $x \in V(\Gamma_r)$ that corresponds to a path of length l in G correspond to those paths of length $l+1$ that have x as a prefix. Γ_r comes together with a natural projection π_r: $V(\Gamma_r) \to V(G)$ which associates to $x \in V(\Gamma_r)$ the endpoint of the path x represents in G. If G is a Δ-regular graph, then $\Gamma_r(G)$ is simply an infinite Δ-regular tree independently of the choice of r. Our basic idea consists of upperbounding the parameter θ of any family of Δ-regular graphs by *the critical probability* for *fractional percolation* on the infinite Δ-regular tree.

1.4 Fractional percolation

Let Γ be an infinite tree together with a privileged vertex r. Classical percolation is concerned with the function $f^1_\Gamma(p)$ defined as the probability that there survives an infinite path rooted at r when we destroy every edge in Γ with probability $1-p$ independently of each other. The associated *critical probability* is $p^1_c(\Gamma) = \inf\{p \mid f^1_\Gamma(p) > 0\}$. For any real number $\alpha, 0 < \alpha \leqslant 1$, let us define the fractional percolation probability $f^\alpha_\Gamma(p)$ as follows. It is the probability that, having independently destroyed every edge in Γ with probability $1-p$, there exists an infinite path $\Pi = [r, x_1, x_2, \dots]$ rooted at r in Γ such that, for every $l > 0$, the segment of Π of length l, $[r, x_1, \dots, x_l]$, contains at least αl surviving edges. This definition is explained in more detail in Section 2. The associated critical probability is defined as $p^\alpha_c(\Gamma) = \inf\{p \mid f^\alpha_\Gamma(p) > 0\}$.

We shall show that, when $\mathscr{G} = (G_m)$ is a family of Δ-regular graphs, then $\theta(\mathscr{G})$ must be upperbounded by $p^{1/2}_c(\Gamma_\Delta)$ where Γ_Δ is the infinite Δ-regular tree. We shall study fractional percolation to make this upperbound explicit: we obtain

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

Theorem 2. *For any* Δ*-regular tree* Γ_Δ,

$$p_c^{1/2}(\Gamma_\Delta) = \frac{1}{2}\left(1 - \sqrt{1 - \frac{1}{(\Delta-1)^2}}\right).$$

2. Fractional percolation and critical probabilities

The results of this section hold for all locally finite infinite trees. By a locally finite infinite tree Γ – simply a tree from now on – we mean a countable connected graph with no cycles, such that each vertex has finitely many neighbours, together with a privileged vertex r called the *root*. We borrow some definitions from Lyons [4]. For any $x \in V(\Gamma)$, denote by $|x|$ the *depth* of x, i.e. the distance between x and r. Write $x \leqslant y$ to mean that vertex x is on the unique path from r to y. Denote by $x \wedge y$ the *join* of x and y, i.e. the vertex of maximum depth that is both on the path from r to x and on that from r to y. A *cutset* K of a tree Γ is a subset of vertices of Γ, such that the connected component of r in $\Gamma \backslash K$ is finite, and such that K is minimal with this property. For any cutset K, denote $|K| = \min_{k \in K} |k|$. A particular cutset is the *sphere of radius* n, $S_n = \{x \in \Gamma, |x| = n\}$. The *branching number* $\operatorname{br}\Gamma$ of Γ is defined as

$$\operatorname{br}\Gamma = \inf\left\{\lambda > 0, \inf_K \sum_{k \in K} \lambda^{-|k|} = 0\right\}$$

where K runs over all cutsets. This concept was introduced in [4], where, among other things, it is proved that for any tree, $p_c^1(\Gamma) = 1/\operatorname{br}\Gamma$.

It is not difficult to check that when Γ is a Δ-regular tree, $\operatorname{br}\Gamma = \Delta - 1$.

Let $\Omega = \{0,1\}^{E(\Gamma)}$. For any $\omega \in \Omega$, denote by $E(\omega)$ the corresponding subset of edges of $E(\Gamma)$. For any $p, 0 < p \leqslant 1$, let μ_p be the measure on $\{0,1\}$ defined by $\mu_p(\{1\}) = p$, $\mu_p(\{0\}) = 1 - p = q$. Define the product probability on Γ, $\mathbf{P}_p = \mu_p^{\otimes E(\Gamma)}$.

Let $\Pi = [x_0, x_1, \ldots, x_m]$ be a finite path in Γ, i.e. $x_i \in V(\Gamma), \{x_i, x_{i+1}\} \in E(\Gamma), x_{i+1} \neq x_{i-1}$. For $\omega \in \Omega$, we shall say that Π is α*-adapted* if each of its segments $[x_0, \ldots, x_l], l = 1 \ldots m$ has at least αl of its edges in $E(\omega)$. We shall say that a vertex x is α*-reachable* if the path from r to x is α-adapted. Denote by $R_\alpha(\omega)$ (or simply $R(\omega)$) the set of α-reachable vertices. We are studying the function

$$f_\Gamma^\alpha(p) = \mathbf{P}_p(\{\omega \in \Omega, R_\alpha(\omega) \text{ is infinite}\}),$$

and its critical probability

$$p_c^\alpha(\Gamma) = \inf\{p \,|\, f_\Gamma^\alpha(p) > 0\}.$$

For $\alpha, p \in]0,1[$, denote $D(\alpha \,||\, p) = \alpha \ln \frac{p}{\alpha} + (1-\alpha)\ln\frac{1-p}{1-\alpha}$. We shall prove:

Theorem 3. *For any tree* Γ,

$$p_c^\alpha(\Gamma) = \sup\left\{p \leqslant \alpha, D(\alpha \,||\, p) < \ln\frac{1}{\operatorname{br}\Gamma}\right\}. \tag{1}$$

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

In particular,

$$p_c^{1/2}(\Gamma) = \frac{1}{2}\left(1 - \sqrt{1 - \frac{1}{\mathrm{br}\,\Gamma^2}}\right). \tag{2}$$

Note that we may always suppose $0 < p \leqslant \alpha$, since it is straightforward to prove that for any tree, $f_\Gamma^\alpha(p) > 0$ when $p > \alpha$ (consider a single infinite path).

We shall prove Theorem 3 in two steps. First we prove the lower bound:

Proposition 1.

$$p_c^\alpha(\Gamma) \geqslant \sup\left\{p \leqslant \alpha, D(\alpha\,||\,p) < \ln\frac{1}{\mathrm{br}\,\Gamma}\right\}.$$

Proof. Fix $\alpha, 0 < \alpha \leqslant 1$.

Let K be any cutset of Γ and define Z_K as the number of vertices of K reachable from the root, i.e.

$$Z_K(\omega) = \sum_{k \in K} \mathbf{1}_{R(\omega)}(k). \tag{3}$$

We have

$$\begin{aligned}\mathbf{E}_p[Z_K] &= \sum_{k \in K} \mathbf{P}_p(k \in R(\omega)) \\ &\leqslant \sum_{k \in K} \sum_{i > \alpha|k|} \binom{|k|}{i} p^i q^{|k|-i}.\end{aligned}$$

Using Stirling's approximation of binomial coefficients,

$$\binom{n}{\lambda n} \leqslant \frac{1}{\sqrt{2\pi n\lambda(1-\lambda)}} \mathrm{e}^{nH(\lambda)} \tag{4}$$

where $H(\lambda) = -\lambda \ln \lambda - (1-\lambda)\ln(1-\lambda)$ denotes the entropy function, we obtain

$$\begin{aligned}\mathbf{E}_p[Z_K] &\leqslant \sum_{k \in K} \sum_{i > \alpha|k|} \exp\left(|k|\left(H\left(\frac{i}{|k|}\right) + \frac{i}{|k|}\ln p + \left(1 - \frac{i}{|k|}\right)\ln(1-p)\right) + o(|k|)\right) \\ &\leqslant \sum_{k \in K} \sum_{i > \alpha|k|} \exp\left(|k| D\left(\frac{i}{|k|}\,||\,p\right) + o(|k|)\right) \\ &\leqslant \sum_{k \in K} \exp\left(|k|\,(D(\alpha\,||\,p) + o(1))\right). \end{aligned} \tag{5}$$

Now then, by the very definition of $\mathrm{br}\,\Gamma$, we know that if $\mathrm{e}^{-D(\alpha||p)} > \mathrm{br}\,\Gamma$, then there exists a sequence K_n of cutsets such that $|K_n| \to \infty$ and the right-hand side of (5) tends to 0 as n goes to $+\infty$. Then, applying Fatou's lemma, this induces that

$$\liminf_{n \to \infty} Z_{K_n}(\omega) = 0 \quad a.s.$$

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

This in turn means that for almost every ω, there exists an index after which $R(\omega)$ is empty, i.e. $R(\omega)$ is almost surely finite. □

To prove the upper bound, we shall actually prove a slightly stronger result. To introduce it we need some more definitions. Let $(u_i)_{i=1}^{\infty}$ be a non-decreasing sequence of positive integers with $\lim_{i\to\infty} u_i = +\infty$. Let us impose, furthermore, that u_i tends to infinity sufficiently slowly, in the sense that $u_{i+1} \leqslant (\sum_{j=1}^{i} u_j)^{1/2}$ for $i \geqslant 1$. Let $(U_i)_{i=1}^{\infty}$ be the associated sequence defined by $U_0 = 0, U_i = \sum_{j=1}^{i} u_j$. Let $\iota(n) = \sup\{i, U_i \leqslant n\}$. For $\omega \in \Omega$, let us say that a finite path $[x_0, x_1, x_2, \dots, x_m]$ is (α, u)-*adapted* if all the subsegments

$$[x_0, \dots, x_{U_1}], [x_{U_1}, \dots, x_{U_2}], \dots, [x_{U_{\iota(m)-1}}, \dots, x_{U_{\iota(m)}}], [x_{U_{\iota(m)}}, \dots, x_m]$$

are α-adapted. For $\omega \in \Omega$, let us say that a vertex x is (α, u)-*reachable* if the path from r to x is (α, u)-adapted. Let $R_{\alpha,u}(\omega)$ (or simply $R_u(\omega)$) denote the set of (α, u)-reachable vertices. We shall now study the probability

$$f_{\Gamma}^{\alpha,u}(p) = \mathbf{P}_p(\omega \in \Omega, R_{\alpha,u}(\omega) \text{ is infinite}).$$

It will turn out that $f_{\Gamma}^{\alpha,u}(p)$ and $f_{\Gamma}^{\alpha}(p)$ have the same critical probabilities. We shall prove:

Proposition 2. *Whenever* $p > \sup\left\{p \leqslant \alpha, D(\alpha\,||\,p) < \ln\dfrac{1}{\mathrm{br}\,\Gamma}\right\}$, *then* $f_{\Gamma}^{\alpha,u}(p) > 0$.

Since, obviously, $f_{\Gamma}^{\alpha,u}(p) \leqslant f_{\Gamma}^{\alpha}(p)$, this last proposition will prove Theorem 3.

Proof of Proposition 2. Let $\Pi_p(m)$ be the probability that a path $[x_0, x_1, \dots, x_m]$ be α-adapted, with the convention that $\Pi_p(0) = 1$. Define

$$f(n) = \left(\prod_{i=1}^{\iota(n)} \Pi_p(u_i)^{-1}\right) \Pi_p(n - U_{\iota(n)})^{-1}.$$

In other words, $1/f(n)$ is the probability that a vertex of depth n is (α, u)-reachable from r.

Next, let ϕ be a *unit flow* on Γ, i.e. a non-negative function on the vertices of Γ such that $\phi(r) = 1$ and $\phi(x) = \sum_{y \succ x, |y|=|x|+1} \phi(y)$ for all $x \in V(\Gamma)$.

Finally, define the random variable

$$X_n^{\phi} = \sum_{s \in S_{U_n}} \phi(s) f(|s|)\, \mathbf{1}_{R_u(\omega)}(s).$$

Let $\mathscr{F}_n$ be the σ-algebra of subsets of Ω generated by the family of events $\{A_x\}_{|x| \leqslant U_n}$ where $A_x = \{\omega \in \Omega, x \in R_u(\omega)\}$. The proof of the proposition consists of:

(i) showing that X_n^{ϕ} is adapted to $\mathscr{F}_n$ and forms a martingale, i.e.

$$\mathbf{E}_p[X_{n+1}^{\phi}\,|\,\mathscr{F}_n] = X_n^{\phi};$$

(ii) showing that whenever $p > \sup\left\{p \leqslant \alpha, D(\alpha\,||\,p) < \ln\dfrac{1}{\mathrm{br}\,\Gamma}\right\}$, then there exists a choice of ϕ such that $\sup_{n \geqslant 1} \mathbf{E}_p[(X_n^{\phi})^2] < \infty$.

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

(iii) arguing that
- since $\{X_n^\phi\}$ is non-negative, it converges a.s. to some X^ϕ.
- whenever $\sup_{n \geqslant 1} \mathbf{E}_p[(X_n^\phi)^2] < \infty$, then $\mathbf{E}_p[X^\phi] = \mathbf{E}_p[X_n^\phi] = 1$ (martingale theory, see [6, Prop. IV.2.7] or [2, p. 24]). Hence, $X^\phi \not\equiv 0$ on a set of non-zero probability, so that, on a set of non-zero probability, $X_n^\phi \not\equiv 0$ for all n, i.e. $\mathbf{P}_p(R_u(\omega)$ is infinite$) > 0$.

Let us prove point (i).

$$\begin{aligned}\mathbf{E}_p[X_{n+1}^\phi \mid \mathscr{F}_n] &= \sum_{x \in S_{U_{n+1}}} \phi(x) f(|x|)\, \mathbf{E}_p[\mathbf{1}_{R_u(\omega)}(x) \mid \mathscr{F}_n] \\ &= \sum_{x \in S_{U_{n+1}}} \phi(x) f(|x|)\, \mathbf{E}_p\left[\mathbf{1}_{R_u(\omega)}(x_{U_n}) \prod_{x_{U_n} \prec y \leqslant x} \mathbf{1}_{R_u(\omega)}(y) \,\middle|\, \mathscr{F}_n\right]\end{aligned}$$

where x_{U_n} denotes the vertex of depth U_n on the path from r to x,

$$\begin{aligned}&= \sum_{x \in S_{U_{n+1}}} \phi(x) f(|x|)\, \mathbf{1}_{R_u(\omega)}(x_{U_n})\, \mathbf{E}_p\left[\prod_{x_{U_n} \prec y \leqslant x} \mathbf{1}_{R_u(\omega)}(y)\right] \\ &= \sum_{x \in S_{U_{n+1}}} \phi(x) f(|x|)\, \Pi_p(u_{n+1})\, \mathbf{1}_{R_u(\omega)}(x_{U_n}) \\ &= \sum_{y \in S_{U_n}} f(|y|)\, \mathbf{1}_{R_u(\omega)}(y) \sum_{x \succ y,\, |x| = U_{n+1}} \phi(x) \\ &= X_n^\phi.\end{aligned}$$

It remains now to prove point (ii). We have

$$\mathbf{E}_p[(X_n^\phi)^2] = \sum_{x,\, y \in S_{U_n}} \phi(x)\, \phi(y) f(U_n)^2\, \mathbf{P}_p(x, y \in R_u(\omega)).$$

Write

$$\mathbf{P}_p(x, y \in R_u(\omega)) \leqslant \left(\prod_{i=1}^{\iota(|x \wedge y|)} \Pi_p(u_i)\right) \Pi_p(u_{\iota(|x \wedge y|)+1}) \left(\prod_{i=\iota(|x \wedge y|)+2}^{n} \Pi_p(u_i)\right)^2$$

to obtain

$$\begin{aligned}\mathbf{E}_p[(X_n^\phi)^2] &\leqslant \sum_{x,\, y \in S_{U_n}} \phi(x)\, \phi(y) f(U_{\iota(|x \wedge y|)+1}) \\ &\leqslant \sum_{|z| \leqslant U_n} f(U_{\iota(|z|)+1}) \sum_{\substack{x \wedge y = z \\ x,\, y \in S_{U_n}}} \phi(x)\, \phi(y) \\ &\leqslant \sum_{|z| \leqslant U_n} f(U_{\iota(|z|)+1}) \sum_{\substack{x,\, y \in S_{U_n} \\ z \leqslant x,\, z \leqslant y}} \phi(x)\, \phi(y)\end{aligned}$$

and since ϕ is a flow, we have

$$\mathbf{E}_p[(X_n^\phi)^2] \leqslant \sum_{|z| \leqslant U_n} \phi(z)^2 f(U_{\iota(|z|)+1}). \tag{6}$$

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

Note that, because u is chosen such that $u_{\iota(n)+1} \leqslant \sqrt{U_{\iota(n)}} \leqslant \sqrt{n}$, $f(U_{\iota(|z|)+1})$ is upperbounded by $f(|z|+\sqrt{|z|})$. The right-hand side of (6) therefore converges when n goes to infinity if

$$\sum_{|z| \leqslant U_n} \phi(z)^2 f(|z|+\sqrt{|z|})$$

converges. Furthermore, we have

$$\Pi_p(m) \geqslant \frac{1}{m}\binom{m}{\alpha m} p^{\alpha m} q^{(1-\alpha)m}.$$

(To check this, use the following combinatorial trick: let $[x_0, x_1, \ldots, x_m = x_0]$ be a cycle with m edges. Notice that any subset of at least αm of its edges must be such that some path $[x_i, x_{i+1}, \ldots, x_m, x_1, \ldots, x_{i-1}]$, $0 \leqslant i \leqslant m-1$, is α-adapted. Hence, when picking each edge with independent probability p, the probability that the resulting subset contains at least αm edges is not more than $m\Pi_p(m)$.)

Hence, using Stirling's approximation of binomial coefficients, as in the proof of Proposition 1,

$$\Pi_p(m) \geqslant \frac{c}{m^{3/2}} \exp(mD(\alpha \,\|\, p))$$

for some constant c. Hence,

$$\begin{aligned} f(n) &\leqslant \prod_{i=1}^{\iota(n)+1} c^{-1} u_i^{3/2} \exp(-u_i D(\alpha\|p)) \\ &\leqslant \exp\left(-(n+\sqrt{n}) D(\alpha \,\|\, p) + \sum_{i=1}^{\iota(n)+1} (\tfrac{3}{2}\ln u_i - \ln c)\right). \end{aligned}$$

Now since $u_i \to +\infty$ as $i \to +\infty$ we must have

$$\frac{\sum_{i=1}^{m} \ln u_i}{\sum_{i=1}^{m} u_i} \to 0 \text{ as } i \to +\infty$$

We deduce from this that, whenever $\mathrm{e}^{-D(\alpha \,\|\, p)}$ is smaller than the radius of convergence $\rho(\phi)$ of the series

$$\sum_{x \in V(\Gamma)} \phi(x)^2 \lambda^{|x|},$$

the sum (6) is bounded independently of n. It is proved in [4] (see Corollary 4.4) that

$$\mathrm{br}\,\Gamma = \sup_{\phi} \rho(\phi)$$

where the supremum is taken over all unit flows of Γ. Hence, whenever $\mathrm{e}^{-D(\alpha \,\|\, p)} < \mathrm{br}\,\Gamma$, there exists some unit flow ϕ such that (6) is bounded independently of n.

Remark. When Γ is a Δ-regular tree, one may simply choose the unit flow, $\phi(x) = (\Delta-1)^{-|x|}$, to obtain the result. □

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

3. Threshold probabilities of cycle codes of graphs

The main result of this section is:

Theorem 4. *Let $\mathscr{G} = (G_m)$ be a sequence of Δ-regular graphs such that $\lim_{m\to\infty} d(G_m) = +\infty$. We have*

$$\theta(\mathscr{G}) \leqslant p_c^{1/2}(\Gamma_\Delta)$$

where Γ_Δ is the Δ-regular tree.

Proof. Fix $p > p_c^{1/2}(\Gamma_\Delta)$. We shall prove the theorem by finding a uniform non-zero lower bound on $f_{G_m}(p)$ for an infinite number of subscripts m.

Without loss of generality, (throw away some G_m) suppose that $d(G_m)$ is an increasing sequence. Let $(u_i)_{i=1}^{\infty}$ be a non-decreasing sequence of positive integers. As in the proof of Proposition 2, let us use the notation $U_0 = 0$, $U_i = \sum_{j=1}^{i} u_j$, $\iota(n) = \sup\{i, U_i \leqslant n\}$.

Let us now impose that the sequence u satisfies

$$u_{\iota(v(G_m))} \leqslant \sqrt{d(G_m)} \tag{7}$$

where $v(G_m)$ denotes the number of vertices in G_m. Equation (7) means that, if we imagine a path $[x_0, \ldots, x_v]$ of length equal to the number of vertices in G, and if we partition the path into segments of non-decreasing lengths u_i, then the largest segment must have a length smaller than the square root of the girth $d(G)$ of G (the choice of the square root is arbitrary, we may replace $\sqrt{n}$ by any function which is $o(n)$). Some thought shows that it is always possible to choose such a sequence u. Let $G_m = G$ be some graph of the sequence $\mathscr{G}$: to lighten notation we omit the subscript m, recalling it only when needed.

Fix some vertex r of G. Consider Γ, the covering tree of G rooted at r. It is a Δ-regular tree with associated root r. The strategy is to compare $f_G(p)$ with $f_\Gamma^{1/2}(p)$. Actually, what we will precisely show is that for any $\epsilon > 0$, there exists m_0 such that for all $m \geqslant m_0$,

$$f_{G_m}(p) \geqslant f_\Gamma^{1/2+\epsilon,\, u}(p). \tag{8}$$

Proposition 2 and Theorem 1 imply that if (8) holds, then $\theta(\mathscr{G}) \leqslant p_c^{1/2+\epsilon}(\Gamma)$ holds. That (8) holds for all ϵ implies the theorem.

Let us therefore prove (8). Let $\alpha = 1/2 + \epsilon$ be such that

$$p > p_c^{\alpha}(\Gamma_\Delta) > p_c^{1/2}(\Gamma_\Delta). \tag{9}$$

Set $\Omega_G = \{0,1\}^{E(G)}$, $\Omega_\Gamma = \{0,1\}^{E(\Gamma)}$. Define the probabilities on G and Γ, respectively, as $\mathbf{P}_p^G = \mu_p^{\otimes E(G)}$ and $\mathbf{P}_p^\Gamma = \mu_p^{\otimes E(\Gamma)}$. For $\omega_g \in \Omega_G$ ($\omega_\gamma \in \Omega_\Gamma$), denote by $E(\omega_g)$ (by $E(\omega_\gamma)$) the corresponding subset of edges of $E(G)$ (of $E(\Gamma)$). Let us denote by $R(\omega_g)$ (by $R(\omega_\gamma)$) the set of vertices 'reachable' from r in G (in Γ). By this we mean the set of vertices $x \in G$ ($x \in \Gamma$), such that there exists a path from r to x in G (in Γ) which is (α, u)-adapted for ω_g (for ω_γ). Finally, to $\omega_g \in \Omega_G$ and $\omega_\gamma \in \Omega_\Gamma$, let us associate two subgraphs $G(\omega_g)$ and $G(\omega_\gamma)$ of G. $G(\omega_g)$ is defined as the subgraph of G induced by $R(\omega_g)$ (the subgraph of G induced by $R \subset V(G)$ is the graph with vertex set R and those edges of $E(G)$ with endpoints in R). $G(\omega_\gamma)$ is defined

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

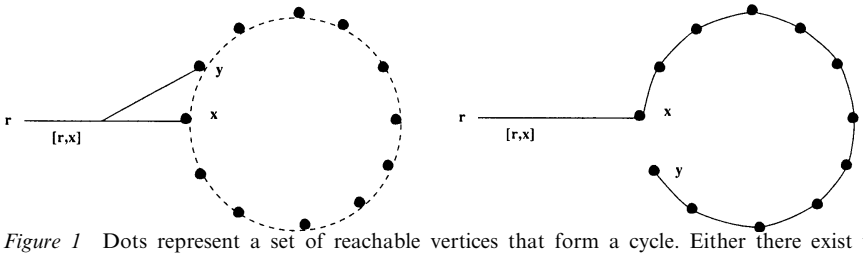


Figure 1 Dots represent a set of reachable vertices that form a cycle. Either there exist two neighbouring vertices x and y with an (α, u)-adapted path $[r, x]$ from r to x and an (α, u)-adapted path from r to y that differs from the concatenation of $[r, x]$ and y, (left picture): or there exists (x, y) such that the (α, u)-adapted path from r to y is as drawn on the right picture.

as the subgraph of G induced by $\pi(R(\omega_\gamma))$, where π is the natural projection from $V(\Gamma)$ to $V(G)$.

Consider now the two events in Ω_G and Ω_γ, respectively:

$$A_G = \{\omega_g \in \Omega_G, G(\omega_g) \text{ contains no cycle}\}$$
$$A_\Gamma = \{\omega_\gamma \in \Omega_\Gamma, G(\omega_\gamma) \text{ contains no cycle}\}.$$

We claim that

- *Claim* 1. $\mathbf{P}_p^G(A_G) = \mathbf{P}_p^\Gamma(A_\Gamma)$.
- *Claim* 2. For m big enough, $f_{G_m}(p) \geqslant 1 - \mathbf{P}_p^{G_m}(A_{G_m})$.

Clearly, if $G(\omega_\gamma)$ has no cycle, then $R(\omega_\gamma)$ must be finite: in other words,

$$\mathbf{P}_p^\Gamma(A_\Gamma) \leqslant 1 - f_\Gamma^{\alpha, u}(p).$$

Therefore, Claims 1 and 2 imply (8).

To prove Claim 1, denote G_R, for $R \subset V(G)$, the subgraph of G induced by R. Define the events of Ω_G and Ω_Γ, respectively,

$$A_G(R) = \{\omega_g \in \Omega_G, R(\omega_g) = R\}$$
$$A_\Gamma(R) = \{\omega_\gamma \in \Omega_\Gamma, R(\omega_\gamma) = R\}.$$

Observe that A_G and A_Γ can be partitioned as

$$A_G = \bigcup_R A_G(R) \quad \text{and} \quad A_\Gamma = \bigcup_R A_\Gamma(R),$$

where unions run over all $R \subset V(G)$ such that G_R contains no cycle, so that

$$\mathbf{P}_p^G(A_G) = \sum_R \mathbf{P}_p^G(A_G(R)) \quad \text{and} \quad \mathbf{P}_p^\Gamma(A_\Gamma) = \sum_R \mathbf{P}_p^\Gamma(A_\Gamma(R)).$$

Because G_R contains no cycle, there is a unique connected subgraph Γ_R of Γ such that $r \in \Gamma_R$ and $\pi(V(\Gamma_R)) = R$. G_R and Γ_R are isomorphic finite trees. Furthermore, since G_R has no cycles, the number I_R^G of edges of G that have exactly one endpoint in R equals the number I_R^Γ of edges of Γ that have exactly one endpoint in $V(\Gamma_R)$. Since the probabilities

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

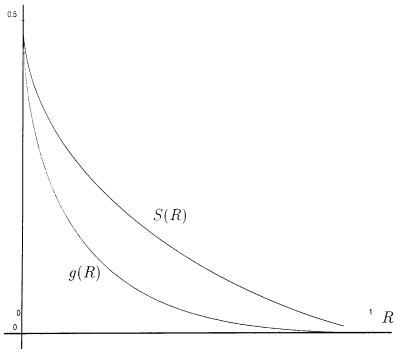


Figure 2

$\mathbf{P}_p^G(A_G(R))$ and $\mathbf{P}_p^\Gamma(A_\Gamma(R))$ are entirely determined by G_R and I_R^G and Γ_R and by I_R^Γ, respectively, we deduce that

$$\mathbf{P}_p^G(A_G(R)) = \mathbf{P}_p^\Gamma(A_\Gamma(R))$$

whenever G_R contains no cycle. Hence Claim 1.

To prove Claim 2, consider $\omega_g \notin A_G$, so that $G(\omega_g)$ contains a cycle. By moving along this cycle one must either

- find a vertex x which is at distance 1 from a vertex y lying on an (α, u)-adapted path from r to x,
- find two adjacent vertices x and y together with an (α, u)-adapted path from r to x not passing through y and an (α, u)-adapted path from r to y not passing through x (see Figure 1).

In both cases, we find a cycle of length $l \geqslant d(G)$ in $G(\omega_g)$ that, by (7), must have at least $\alpha l - 2\sqrt{l}$ edges in $E(\omega_g)$. Because α is chosen to satisfy (9), this is more than $l/2$ when the girth $d(G)$ is big enough. □

4. Concluding remarks

Since the asymptotic rate of large Δ-regular graphs is $R = 1 - 2/\Delta$, Theorem 4 gives an asymptotic upper bound $g(R)$ on the maximum channel probability p for which cycle codes of graphs can achieve vanishing decoding error probability. Figure 2 compares $g(R)$ to Shannon's upper bound $S(R)$ for codes in general:

$$g(R) = \frac{1}{2}\frac{(1-\sqrt{R})^2}{1+R}$$

$$S(R) = H_2^{-1}(1-R)$$

where $H_2(x) = \dfrac{1}{\ln 2} H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function.

The class of cycle codes of Δ-regular graphs actually achieves the bound $g(R)$.

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press

J.-P. Tillich has found this to be true for some classes of Ramanujan graphs. This is the object of a forthcoming paper [7].

References

[1] Berge, C. (1970) *Graphes et hypergraphes*. Dunod.
[2] Dellacherie, C. and Meyer, P. A. (1980) *Probabilités et Potentiel, Tome II*. Hermann.
[3] Hakimi, S. L. and Bredeson, J. G. (1968) Graph theoretic error-correcting codes. *IEEE Trans. Inf. Theory* **IT-14** 584–591.
[4] Lyons, R. (1990) Random walks and percolation on trees. *Annals of Probability* **18** 931–958.
[5] Margulis, G. (1974) Probabilistic characteristics of graphs with large connectivity. *Problemy Peredachi Informatsii* **10** 101–108.
[6] Neveu, J. (1972) *Martingales à Temps Discret*. Masson.
[7] Tillich, J.-P. and Zémor, G. Optimal cycle codes constructed from Ramanujan graphs. *SIAM J. on Discrete Math*. To appear.
[8] Zémor, G. (1994) Threshold effects in codes. In *Algebraic Coding: Lecture Notes in Computer Science 781*. Springer-Verlag, pp. 278–286.

https://doi.org/10.1017/S0963548396002878 Published online by Cambridge University Press