On the Local Distinguishing Numbers of Cycles

C. T. Cheng L. J. Cowen

Department of Mathematical Sciences
Johns Hopkins University
Baltimore, MD 21218
cheng,cowen@mts.jhu.edu

September 2, 1998

Abstract

Consider the induced subgraph of a labeled graph G rooted at vertex v, denoted by N_v^i, where $V(N_v^i) = \{u : 0 \leq d(u,v) \leq i\}$. A labeling of the vertices of G, $\Phi : V(G) \to \{1, ..., r\}$ is said to be i-local distinguishing if $\forall u, v \in V(G), u \neq v$, N_v^i is not isomorphic to N_u^i under Φ. The ith local distinguishing number of G, $\mathrm{LD}^i(G)$ is the minimum r such that G has an i-local distinguishing labeling that uses r colors. $\mathrm{LD}^i(G)$ is a generalization of the distinguishing number $\mathrm{D}(G)$ as defined in [1].

An exact value for $\mathrm{LD}^1(C_n)$ is computed for each n. It is shown that $\mathrm{LD}^i(C_n) = \Omega(n^{\frac{1}{2i+1}})$. In addition, $\mathrm{LD}^i(C_n) \leq 24(2i+1)n^{\frac{1}{2i+1}}(\log n)^{\frac{2i}{2i+1}}$ for constant i was proven using probabilistic methods. Finally, it is noted that for almost all graphs G, $\mathrm{LD}^1(G) = O(\log n)$.

1 Introduction

The following problem was recently reintroduced by Albertson and Collins [1]. Suppose a professor has a set of n keys on a circular key ring that look similar enough to each other so that they are indistinguishable to the naked eye. To tell the keys apart, he attaches a colored marker on each key. How many different colors of markers must he have, in order to label the keys, so that he can distinguish the keys from each other? When $n \geq 6$, two colors suffice. The professor simply chooses 5 contiguous keys, labels them with colors $1, 2, 1, 2$ and 2 respectively, and labels the rest of the keys with color 1. Since the string 12122 is not a palindrome, i.e., it is not the same as its reverse, the professor can identify each key by its clockwise distance from the two contiguous keys colored 2. However, perhaps surprisingly, it can be shown that when $n = 3, 4, 5$, three different colors are required.

If n is large, notice that the professor has to look all the way across the key ring and count nearly up to $n/2$ keys to determine his key. Going back to our original problem, suppose we add the restriction that in order to determine the key he is holding, the professor is allowed to look only at that key and at most i keys to the left and right of that key. Now, what is the minimum number of colors he needs? This number, which we call the *ith local* *distinguishing number of the cycle*, is the main subject of this paper. But we can define the ith local distinguishing number in a more general setting as follows.

The answer to the original problem and the new problem are both dependent on the fact that the keyholder was circular. If, instead, the keys were suspended from a straight rod, for example, the answer to the original problem would change: it is not hard to see that two colors suffice for all $n \geq 2$. In [1], Albertson and Collins generalized the original problem to arbitrary graphs. Given a graph G, they defined the *distinguishing number of G,* denoted $D(G)$, to be the minimum number of colors so that there exists a coloring of G that uses this number of colors whose group of color-preserving automorphisms is trivial. In particular, the group of automorphisms of the uncolored cycle consists of all rotations of the cycle and flips about each vertex of the cycle. However, for $n \geq 6$, the only automorphism of an n-vertex cycle when 2-colored as we described in the first paragraph is the identity since the colors of the vertices must be preserved. In particular, the 5 contiguous vertices labeled $1, 2, 1, 2, 2$ must be mapped to themselves. Thus, $D(C_n) = 2$ if $n \geq 6$.

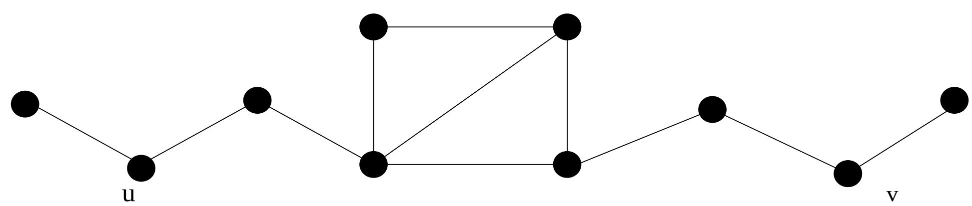

Figure 1: $N_u^2 \cong N_v^2$ and thus, $\mathrm{LD}^2(G) > 1$. By assigning a color to v different from those of the remaining vertices, we can easily prove that $\mathrm{LD}^2(G) = 2$.

Let $G = (V(G), E(G))$ be a graph and $v \in V(G)$. Let N_v^i be the neighborhood of v out to distance i in G; that is, N_v^i is the induced subgraph of G rooted at v for which $V(N_v^i) := \{u : 0 \leq d(v,u) \leq i\}$. If G is a colored graph, we also refer to N_v^i as the *ith naming subgraph of* v. The ith naming subgraph of u and the ith naming subgraph of v are said to be *isomorphic* if and only if there is an isomorphism from N_u^i to N_v^i that maps u onto v, and that additionally preserves colors. A *labeling,* or *coloring,* of the vertices of G, $\Phi : V(G) \rightarrow \{1, ...r\}$, is said to be *i-locally distinguishing* if no two vertices have isomorphic ith naming subgraphs. Consequently, we say that Φ is an LD^{i}-labeling of G. The ith local distinguishing number of G, denoted by $\mathrm{LD}^i(G)$, is the minimum r such that G has an LD^{i} labeling that uses r colors. (See figure 1 for an example.) We note that $\mathrm{D}(G)$ is a lower bound on $\mathrm{LD}^i(G)$, for all i.

This paper considers $\mathrm{LD}^i(C_n)$. When $i = 0$, $\mathrm{LD}^0(C_n)$ is clearly n, and $\mathrm{LD}^i(C_n)$, for fixed constant i, will clearly tend to infinity as $n \rightarrow \infty$. It is not clear that $\mathrm{LD}^i(C_n)$ will be strictly non-decreasing, however, as a function of n. In Section 2, we solve $\mathrm{LD}^1(C_n)$ exactly, and prove that it increases monotonically as a function of n. The information- theoretic lower bound we derived in the preliminary section implies that $\mathrm{LD}^i(C_n) = \Omega(n^{\frac{1}{2i+1}})$. In Section 3, we use probabilistic arguments to give upper bounds for $\mathrm{LD}^i(C_n)$ when i is a constant. In particular, we show $LD^i(C_n) \leq 24(2i+1)n^{\frac{1}{2i+1}}(\log n)^{\frac{2i}{2i+1}}$.

When $i = \text{diameter}(G)$, $\mathrm{LD}^i(G) = \mathrm{D}(G)$. Thus, $\mathrm{LD}^{\lfloor \frac{n}{2} \rfloor}(C_n) = \mathrm{D}(C_n)$. In Section 4, we show that, in fact, for $i = \lceil \log_2 n + 1 \rceil$, $\mathrm{LD}^i(C_n) = \mathrm{D}(C_n)$. This implies that, in our original problem, there exists a labeling of the keys with two colors such that the professor can always identify the key he is holding by searching at most $O(\log n)$ neighboring keys on both sides, instead of searching the entire key ring.

At this point, we know little about $\mathrm{LD}^i(G)$ for graphs other than cycles. However as a first step, we show that for almost all graphs on n vertices, when $n >> 0$, $\mathrm{LD}^1(G) = O(\log n)$ in Section 5.

Preliminaries. Consider a vertex v in C_n. When $n > 2i+1$, N_v^i is a path with $2i+1$ vertices centered at vertex v. Let this path consist of vertices $(v_{-i}, v_{-(i-1)}, \ldots, v_{-1}, v_0, v_1, \ldots, v_{i-1}, v_i)$ where v_j and and v_{-j} are the two vertices at distance j from $v_0 = v$.

Fix the reading of labels in C_n in one direction. Let Φ be a general labeling of C_n. We associate a $(2i+1)$-tuple to v, namely,

$$\mathrm{L}(v) := (\Phi(v_{-i}), \Phi(v_{-(i-1)}), ..., \Phi(v_{-1}), \Phi(v_0), \Phi(v_1), ...\Phi(v_{i-1}), \Phi(v_i)).$$

The reversed-tuple we denote by

$$\overline{\mathrm{L}}(v) := (\Phi(v_i), \Phi(v_{i-1}), ..., \Phi(v_1), \Phi(v_0), \Phi(v_{-1}), ...\Phi(v_{-(i-1)}), \Phi(v_{-i})).$$

We say u *is equivalent to* v and use the notation $u \simeq v$ when either

(i) $u \simeq v$ via a *direct match* where $\mathrm{L}(u) = \mathrm{L}(v)$, or

(ii) $u \simeq v$ via a *flip* where $\mathrm{L}(u) = \overline{\mathrm{L}}(v)$.

If $u \simeq v$ either by a flip or a direct match then they have isomorphic naming subgraphs in C_n. That is, in an LD^{i}-labeling of C_n, no two vertices u and v have $u \simeq v$.

We close this section with a simple information-theoretic lower bound on $\mathrm{LD}^i(C_n)$ based on the number of inequivalent $2i+1$-tuples that can be formed using r colors.

Lemma 1.1. *If* $n > \frac{r^{i+1}(r^i+1)}{2}$, *then* $\mathrm{LD}^i(C_n) > r$.

Proof: There are r^{2i+1} $(2i+1)$-tuples that can be generated using r colors. Of these, r^{i+1} are palindromes, i.e. $(a_1, a_2, \ldots, a_{2i}, a_{2i+1}) = (a_{2i+1}, a_{2i}, \ldots, a_2, a_1)$. The rest are asymmetric and their flips also appear in the enumeration of the $(2i+1)$-tuples. Thus, the total number of inequivalent $(2i+1)$-tuples is $\frac{r^{2i+1}-r^{i+1}}{2} + r^{i+1} = r^{i+1}(\frac{r^i+1}{2})$. If $n > r^{i+1}(\frac{r^i+1}{2})$, any r-labeling of C_n would map at least two vertices of C_n to the same $(2i+1)$-tuple. The lower bound on $\mathrm{LD}^i(C_n)$ follows. □

2 On the 1-local distinguishing number of the cycle

When $i = 1$, N_v^1 consists of v and its two neighbors. In this section, $\mathrm{LD}^1(C_n)$ is solved exactly using a constructive proof.

First, we ask a related question: what is the largest cycle that can be 1-locally distinguished using r colors? Define $t(r)$ as the number of inequivalent triples that can be generated using r colors. From the proof of Lemma 1.1, $t(r) = r^2(\frac{r+1}{2})$. Clearly, $C_{t(r)}$ is the largest possible cycle that can be labeled with r colors so that the labeling is 1-locally distinguishing. Furthermore, whenever $n > t(r)$, $\mathrm{LD}^1(C_n) > r$.

Lemma 2.1. *(i) If r is odd,* $\mathrm{LD}^1(C_{t(r)}) = r$.
(ii) If r is even, $\mathrm{LD}^1(C_{t(r)-r}) = r$ *and* $\mathrm{LD}^{(1)}(C_{t(r)-r+j}) > r$ *for* $j > 0$, $j \in \mathbf{Z}$.

Proof: As in the previous section, we say two triples are equivalent if they are equivalent via a direct match or a flip, i.e. $(s,t,u) \simeq (x,y,z)$ if $s = x, t = y, u = z$ or $s = z, t = y, u = x$.

Instead of labeling the vertices of C_n directly, we describe a tour on all $n = t(r)$ inequivalent triples such that whenever the tour traverses the edge from (s,t,u) to (x,y,z) then $t = x$ and $u = y$. We call this property the *contiguity constraint*. If the contiguity constraint is maintained, then n contiguous triples represent the naming subgraphs of the n contiguous vertices in C_n. The labeling of C_n follows naturally, as shown in figure 2.

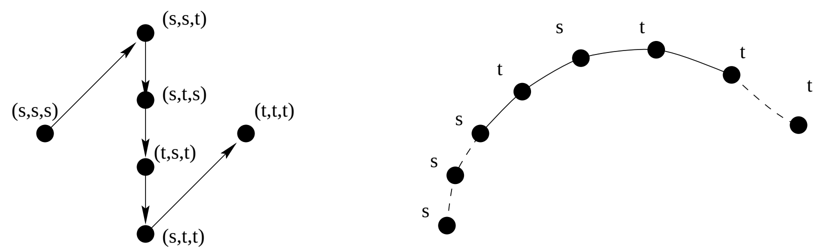

Figure 2: The $s-t$ path and the corresponding labeling in C_n

Let s and t be any two distinct labels so that $s \neq t$. Then there exists a path from (s,s,s) to (t,t,t) which traverses all inequivalent triples that uses labels s and t only. In particular, we consider the path in figure 2 and call it the $s-t$ path.

When r is odd, the complete graph on r vertices K_r is Eulerian. Let E_r be an Euler tour on K_r. We construct the tour on $C_{t(r)}$ based on E_r as follows: whenever the edge (s,t) is traversed on E_r, the $s-t$ path is traversed on the tour. If (s,s,s) has been traversed before, skip to (s,s,t). Since K_r is Eulerian whenever r is odd, all triples which use only two distinct numbers must be traversed by this tour.

Consider all triples (s,t,u) where s,t and u are all distinct. There is a path that traverses the three inequivalent triples involving s,t, and u as shown in figure 3. We insert the $s-t-u$ path into the $s-t$ path as shown in figure 4. Notice that the three vertices could have been inserted in the $t-u$ path or $u-s$ path. We have now traversed all $t(r)$ triples.

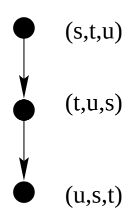

Figure 3: The $s-t-u$ path.

When r is even, not all $t(r)$ triples can be traversed in a proper tour. This is so because whenever (u,s,s) is traversed, (s,s,t), $u \neq t$, must be traversed directly after it or after the tour goes through (s,s,s). Thus, for a fixed s, only an even number of vertices of the form (s,s,t), $s \neq t$ can be traversed. When r is even, there is an odd number of vertices of the form (s,s,t), $s \neq t$. Hence, at least r triples *must* be skipped in touring the triples. So $\mathrm{LD}^1(C_{t(r)-r+j}) > r$ for $j > 0$.

When r is even, K_r is not Eulerian. Delete a maximum matching in K_r so all the vertices have even degrees. Note that $\frac{r}{2}$ edges were deleted. Call this new graph K'_r. As described

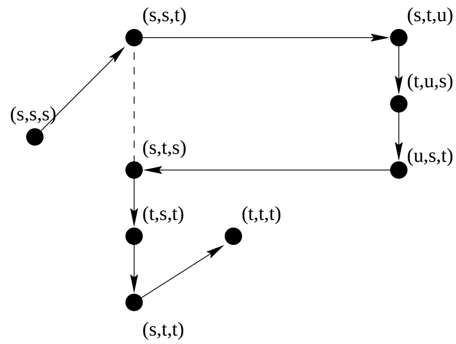

Figure 4: Extending the $s-t$ path.

above, construct a tour on the triples based on the Euler tour of K_r'. Now, the triples of the form (s,t,u) can be inserted in the tour since only one of the edges (s,t), (s,u) was deleted from K_r. However, there are triples that are missing in the tour. These are the triples that use only two distinct labels s,t such that (s,t) was one of the edges deleted from K_r. Insert the path (s,t,s), (t,s,t) after the triple (u,s,t) is visited. We have skipped exactly 2 triples per (s,t) pair: (s,s,t) and (s,t,t); and since $\frac{r}{2}$ edges were deleted, exactly r triples were skipped in this tour. $\square$

We are now ready for the main result of this section. The previous lemma showed that if n is odd and is exactly equal to $t(r)$, the number of inequivalent triples that can be generated from r colors, there was a labeling of C_n based on a tour that visited all these triples. We, then, concluded that $\mathrm{LD}^1(C_n) = r$. We also had a similar result when n is even and equal to $t(r) - r$. In order to handle cycle lengths between $n = t(r)$ and $n = t(r+1)$, we now show how to remove pieces of these tours to construct shorter tours that still fit together.

Theorem 2.2. *Given C_n, let $k \in \mathbf{R}$ s.t. $n = \frac{k^2(k+1)}{2}$. Let $r = \lceil k \rceil > 2$.*
(i) If r is odd, $\mathrm{LD}^1(C_n) = r$.
(ii) If r is even and $n \le \frac{r^2(r+1)}{2} - r$, then $\mathrm{LD}^1(C_n) = r$; otherwise, $\mathrm{LD}^1(C_n) = r+1$.

Proof: From Lemma 1.1, C_n needs at least r colors to have a 1-locally distinguishing labeling. From Lemma 2.1, when r is even and $n > \frac{r^2(r+1)}{2} - r$, C_n needs at least $r+1$ colors.

Our strategy now is to modify the $C_{t(r)}$ and $C_{t(r)-r}$ tours we have constructed in Lemma 2.1. We remove paths to obtain smaller tours which still maintain the contiguity property. The following paths and vertices were traversed in $C_{t(r)}$ when r was odd and in $C_{t(r)-r}$ when r was even:

- r paths of length 0 that go through (s,s,s). We call these paths TYPE 1.
- $\binom{r}{2}$ paths of length 1 that go through (s,t,s), (t,s,t). We call these paths TYPE 2.
- $\binom{r}{3}$ paths of length 2 that go through (s,t,u), (t,u,s), (u,s,t). We call these paths TYPE 3.

Notice that it is possible to skip the above paths in the tour by connecting the two neighbors at both ends of the path and the contiguity property is still maintained in the new tour. See figure 5.

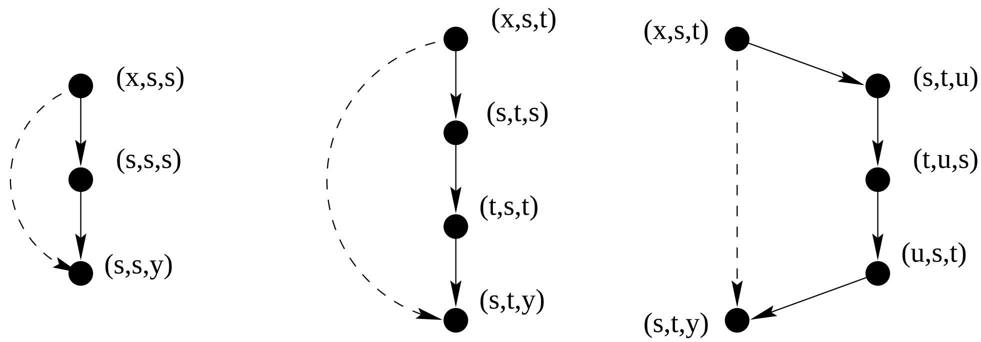

Figure 5: Short-cutting the tours in Lemma 2.1 while maintaining the contiguity constraint.

Let $r \geq 3$ be odd, we shall show that there exists an r-labeling for C_n whenever $t(r-1)-(r-1) < n \leq t(r)$. Denote the tour that goes through all $t(r)$ triples as T.

To obtain a tour on n triples:

(i) when $n = t(r) - 2j$, $0 \leq j \leq \binom{r}{2}$, skip j TYPE 2-paths in T.

(ii) when $n = t(r) - 1 - 2j$, $0 \leq j \leq \binom{r}{2}$, skip a TYPE 1-path and j TYPE 2-paths in T. Denote the tour that goes through $t(r) - 2\binom{r}{2}$ triples as T'.

(iii) when $n = t(r) - 2\binom{r}{2} - 3z$, $0 \leq z \leq \binom{r}{3}$, start with T' and skip z TYPE 3-paths..

(iv) when $n = t(r) - 2\binom{r}{2} - 1 - 3z$, $0 \leq z \leq \binom{r}{3}$, start with T', skip another TYPE 1-path and z TYPE 3-paths.

(v) when $n = t(r) - 2\binom{r}{2} - 2 - 3z$, $0 \leq z \leq \binom{r}{3}$, start with T', skip two TYPE 1-paths and skip z TYPE 3-paths.

Hence, when $t(r) - 2\binom{r}{2} - 3\binom{r}{3} = r^2 \leq n \leq t(r)$, $\mathrm{LD}^i(C_n) \leq r$. But $t(r-1)-(r-1) > r^2$ when $r \geq 5$. When $r = 3$, it is easy to check that a 3-labeling exists for C_n when $7 \leq n \leq 9$. Thus, claim (i) and (iib) follow.

When $r \geq 4$, r even, we use the same technique above to show that an r-labeling for C_n exists whenever $t(r-1) < n \leq t(r) - r$. We reiterate that the r triples skipped in constructing T for the labeling of $C_{t(r)-r}$ were not part of the paths skipped above. Furthermore, $t(r) - r - 2\binom{r}{2} - 3\binom{r}{3} = r^2 - r < t(r-1)$ for $r \geq 3$.

Finally, we note that when $r = 2$, $2 \leq n \leq 6$. In these cases, $\mathrm{LD}^1(C_n) = 3$. $\square$

For $j = 1, 2$, if $n_j = \frac{k_j^2(k_j+1)}{2}$ let $r_j = \lceil k_j \rceil$. If $n_1 < n_2$ then $r_1 \leq r_2$. Theorem 2.2 implies that $\mathrm{LD}^1(C_{n_1}) \leq \mathrm{LD}^1(C_{n_2})$. That is, $\mathrm{LD}^1(C_n)$ is monotonic for cycles.

3 An Upper Bound on $\mathrm{LD}^i(C_n)$

The lower bounds from Lemma 1.1 imply $\mathrm{LD}^i(C_n) = \Omega(n^{\frac{1}{2i+1}})$. In this section, we give an upper bound on $\mathrm{LD}^i(C_n)$ for constant i that is an $O(\log n)$ factor off the lower bound. The proof uses probabilistic methods, but we remark that a standard argument where we color at random with sufficient number of colors so that with high probability, no pair of vertices have isomorphic naming subgraphs yields a poor upper bound on the number of colors needed for small i. Instead, we use a two-stage coloring procedure. First, we color with a smaller number of colors than would be needed to distinguish all i-th naming subgraphs. However, the number of colors we used is big enough so that the size of any class of vertices with equivalent i-naming subgraphs is $O(\log n)$. In the second stage, we then greedily refine the coloring so that vertices that belong to the same equivalence class according to the stage 1-coloring have non-isomorphic naming subgraphs after the refinement.

We prove that there exists a coloring with $O(n^{\frac{1}{2i+1}}(\log n)^{\frac{2i}{2i+1}})$ colors, assigning each vertex an ordered pair of colors (x, y) where x is chosen from a set of $(n/\log n)^{\frac{1}{2i+1}}$ colors, and y is chosen from a set of $O(\log n)$ colors.

Fix i. For a coloring Φ of the vertices of C_n, define $S_v = \{u : u \simeq v \text{ under } \Phi\}$. We will first show that a coloring of C_n exists where the maximum size of S_v, for any v, is not too large.

Lemma 3.1. *There exists a coloring of C_n with $(n/\log n)^{\frac{1}{2i+1}}$ colors such that* $\max_v |S_v| = 24(2i+1)\log n$.

Proof: Color the vertices of C_n randomly with $(n/\log n)^{\frac{1}{2i+1}}$ colors, i.e. for each vertex in C_n select its color uniformly at random from the set $\{1, 2, ..., (n/\log n)^{\frac{1}{2i+1}}\}$. We show the probability that this coloring has greater than $O(\log n)$ vertices with isomorphic ith naming subgraph is less than 1. Thus, the desired coloring must exist.

Let J_k be the set that contains the pair (u, v) such that N_u^i and N_v^i overlap in k positions. Let A_{uv} be the event that $u \simeq v$ via a direct match and B_{uv} be the event that $u \simeq v$ via a flip. For simplicity, let r be the number of colors used for the labeling of C_n.

Claim 3.2. *Suppose we select the color of each vertex in C_n uniformly at random from the set $\{1, 2, \ldots, r\}$. Then*

$$P(u \simeq v) = \begin{cases} \frac{2}{r^{2i+1}} & \text{if } (u,v) \in J_0 \\ \frac{1}{r^{2i+1}} + \frac{1}{r^{2i+1-\lceil \frac{k}{2} \rceil}} & (u,v) \in J_k, 1 \le k \le 2i. \end{cases}$$

Proof of claim: If $(u, v) \in J_0$ then the labels of $N_u^{(i)}$ and $N_v^{(i)}$ are independent. It follows that $P[A_{uv}] = \frac{1}{r^{2i+1}}$.

Otherwise, suppose $(u, v) \in J_k, 0 < k \le 2i$. Without loss of generality, let $\Phi(v_{-i}) = \Phi(u_{i-k+1}), \Phi(v_{-i+1}) = \Phi(u_{i-k+2}), ..., \Phi(v_{-i+k-1}) = \Phi(u_i)$. Thus, the labels of $u_{-i}, u_{-i+1}, ..., u_{i-k}$ are sufficient to determine all the labels in N_u^i and N_v^i if the event A_{uv} is to occur. Hence,

$$P[A_{uv}] = \frac{r^{2i+1-k}}{r^{2(2i+1)-k}} = \frac{1}{r^{2i+1}}.$$

If $(u, v) \in J_0$, $\mathrm{P}[B_{uv}] = \frac{1}{r^{2i+1}}$. Suppose N_u^i and N_v^i overlap in k positions. Consider the graph $N_{uv} = N_u^i \cup N_v^i$. N_{uv} has $2(2i+1) - k$ vertices. It is not difficult to see that the labels of N_{uv} must form a palindrome for B_{uv} to occur. Thus,

$$P[B_{uv}] = \frac{r^{2i+1-\lfloor \frac{k}{2} \rfloor}}{r^{2(2i+1)-k}} = \frac{1}{r^{2i+1-\lceil \frac{k}{2} \rceil}}.$$

$\square$

It is interesting to note that the above claim implies it is more likely for $u \simeq v$ when their naming subgraphs overlap.

To bound the size of the equivalence classes S_v, we do not consider all N_u^i at once because the naming subgraphs may overlap and the colorings would not be independent. Instead, we partition the vertices into $2i+1$ sets each with non-overlapping naming subgraphs as follows. Fix a vertex v, and renumber the vertices $v = 0, 1, 2, ..., n-1$ clockwise around the cycle from v. Let I_{uv} be the indicator random variable for the event $A_{uv} \cup B_{uv}$. Partition the vertices of C_n into $2(2i+1)$ sets $A_0, A_1, ..., A_{4i+1}$ so that if $0 < u \leq \lceil \frac{n}{2} \rceil$ and $j \equiv u \pmod{2i+1}$ then $u \in A_j$. If $u > \lceil \frac{n}{2} \rceil$ and $j \equiv u \pmod{2i+1}$ then $u \in A_{2i+1+j}$. Clearly, the naming subgraphs of any two vertices in A_j, for any j, do not overlap by the way the partitions were constructed and $|A_j| \leq \lceil \frac{n}{2(2i+1)} \rceil$. Furthermore each set A_j has at most one vertex whose naming subgraphs overlap with N_v^i. Let $\delta = 12(2i+1) - 1 = 24i + 11$. Then,

$$
\begin{aligned}
P[|S_v| > 2(1+\delta)\log n] &= P[\sum_{v \in V(C_n)} I_{uv} > 2(1+\delta)\log n] \\
&\leq \sum_{j=0}^{4i+1} P[\sum_{u \in A_j} I_{uv} > 2(1+\delta)\frac{\log n}{2(2i+1)}], \qquad (1)
\end{aligned}
$$

since the A_j's form a partition of $V(C_n)$. From Claim 3.2, it follows that $\frac{\log n}{2(2i+1)} \leq E[\sum_{u \in A_j} I_{uv}] \leq \frac{2 \log n}{2i+1}$. Hence,

$$
\begin{aligned}
P[\sum_{u \in A_j} I_{uv} > 2(1+\delta)\frac{\log n}{2(2i+1)}] &\leq P[\sum_{u \in A_j} I_{uv} > \frac{1+\delta}{2} E[\sum_{u \in A_j} I_{uv}]] \\
&\leq 2^{-\frac{1+\delta}{2} E[\sum_{u \in A_j} I_{uv}]} \qquad (2) \\
&\leq 2^{-\frac{(1+\delta)}{2} \frac{\log n}{2(2i+1)}} \\
&= n^{-\frac{1+\delta}{4(2i+1)}} \\
&= n^{-3}
\end{aligned}
$$

where (2) follows from the fact that if $u, w \in A_j$ then u and w have disjoint naming subgraphs. The events I_{uv} and I_{uw} must be independent from each other and the Chernoff's bounds (see e.g. [6], p.72) can then be applied to these events. Substituting the above result for δ on the right hand side of (1),

$$
P[|S_v| > 2(1+\delta)\log n] \leq \sum_{j=0}^{4i+1} \frac{1}{n^3} = \frac{4i+1}{n^3}
$$

and summing over all v,

$$
P[|S_v| > 2(1+\delta)\log n, \text{ for some } v] \leq \sum_{v \in V(C_n)} P[S_v > 2(1+\delta)\log n] \leq \frac{4i+1}{n^2}.
$$

$\square$

We note that in fact we have found a coloring with high probability and since the condition (that the coloring should induce only small equivalence classes) is verifiable in linear time, this implies a simple randomized algorithm to find this coloring. To prove only the existence of such a coloring, we could have reduced the number of colors in this stage by a constant factor.

Theorem 3.3. *For any constant i, $LD^i(C_n) \leq 24(2i+1)n^{\frac{1}{2i+1}}(\log n)^{\frac{2i}{2i+1}}$.*

Proof: Label C_n such that for any vertex v, $|S_v| < 24(2i+1)\log n$. Such a coloring is guaranteed to exist by Lemma 3.1. Let this labeling be Φ_1. Now consider the vertices in clockwise order around the cycle, starting arbitrarily. When we reach v, we say v has been *visited* and re-color v with the color label $\Phi(v) = (\Phi_1(v), \Phi_2(v))$ where $\Phi_1(v)$ is its label under the old coloring and $\Phi_2(v)$ is chosen greedily from the set $K = \{1, 2, .., 24(2i+1)\log n\}$ as follows: choose $\Phi_2(v)$ to be the first color in the set K which *does not* appear in the set $\{\Phi_2(u) | u \text{ is visited }, u \simeq v \text{ under } \Phi_1\}$. Such a new color always exists by the maximum size of the equivalence classes in the coloring Φ_1. Now $\Phi(u) = \Phi(v)$ if and only if $\Phi_1(u) = \Phi_1(v)$ and $\Phi_2(u) = \Phi_2(v)$. But we have chosen $\Phi_2(u) \neq \Phi_2(v)$ whenever $\Phi_1(u) = \Phi_1(v)$. Thus, Φ is a 1-distinguishing labeling that uses at most $24(2i+1)n^{\frac{1}{2i+1}}(\log n)^{\frac{2i}{2i+1}}$ colors. □

4 Looking out log

We know that $\mathrm{LD}^{\lfloor \frac{n}{2} \rfloor}(C_n) = 3$ when $n = 3, 4, 5$ and $\mathrm{LD}^{\lfloor \frac{n}{2} \rfloor}(C_n) = 2$ when $n > 5$ since $\mathrm{LD}^{\lfloor \frac{n}{2} \rfloor}(C_n) = \mathrm{D}(C_n)$. On the other hand, for $\mathrm{LD}^i(C_n) = 2$, $i = \Omega(\log n)$ by the information theoretic lower bound (Lemma 1.1). Here we show $\mathrm{LD}^{(\lceil \log n+1 \rceil)}(C_n) = 2$ for $n > 5$.

Theorem 4.1. *$LD^{(\lceil \log n+1 \rceil)}(C_n) = D(C_n)$.*

Proof: Let v be a vertex in C_n. When $2 \leq n \leq 11$, $N_v^{\lceil \log n+1 \rceil}$ includes all vertices in C_n. Thus, $\mathrm{LD}^{\lceil \log n+1 \rceil}(C_n) = D(C_n)$ trivially. So assume $n > 11$. We again use a probabilistic argument. This time a straightforward random labeling of C_n is sufficient to show our result. For simplicity, we assume $n = 2^j, j$ a positive integer.

Let Φ be a uniform random labeling of the vertices in C_n with two colors. We shall show with positive probability, Φ is an $\mathrm{LD}^{\lceil \log n+1 \rceil}$ labeling. We keep the notation from Section 3. Note we never consider events where vertex v is compared to itself.

$$\begin{aligned} P(\Phi \text{ is a bad labeling}) &= P(\bigcup_{(u,v) \in VxV} A_{uv} \cup \bigcup_{(u,v) \in VxV} B_{uv}) \\ &\leq P(\bigcup_{(u,v) \in VxV} A_{uv}) + P(\bigcup_{(u,v) \in VxV} B_{uv}). \end{aligned}$$

We solve for $\mathrm{P}[\bigcup_{(u,v) \in VxV} A_{uv}]$ and $\mathrm{P}[\bigcup_{(u,v) \in VxV} B_{uv}]$ separately. From Lemma 3.2, $P[A_{uv}] = \frac{1}{2^{2\log n+3}} = \frac{1}{8n^2}$. So,

$$P[\bigcup_{(u,v)\in VxV} A_{uv}] \leq \binom{n}{2}\frac{1}{8n^2} \leq \frac{1}{16}.$$

Let $(u,v) \in J_0$. From Lemma 3.2, $\mathrm{P}[B_{uv}] = \frac{1}{8n^2}$. If we fix u, there are $n - (4\log n + 6) + 1$ vertices whose naming subgraphs do not overlap with u. Thus, $\sum_{(u,v)\in J_0} P[B_{uv}] = \frac{1}{8n^2}(\frac{n(n-4\log n-5)}{2}) \leq \frac{1}{16}$.

Let $(u,v) \in J_k$, $1 \leq k \leq 2i$, N_{uv} has $2(2\log n + 3) - k$ vertices and when B_{uv} occurs, the labels of N_{uv} form a palindrome . If we remove the same number of vertices from the endpoints of N_{uv}, then the labels still form a palindrome. Hence, there must be two vertices i and j where N_{ij} is a palindrome embedded in N_{uv} and $N_i^{(1)}$ and $N_j^{(1)}$ overlap in $2\log n + 1$ or $2\log n + 2$ positions depending on the parity of k. This implies

$$\bigcup_{(u,v)\in J_k, k>0} B_{uv} \subseteq \bigcup_{(u,v)\in J_{2\log n+1}} B_{uv} \cup \bigcup_{(u,v)\in J_{2\log n+2}} B_{uv}.$$

Now, when $(u,v) \in J_k$ for $k = 2\log n + 1$ or $k = 2\log n + 2$, $P[B_{uv}] = \frac{1}{4n}$. If we fix u, there are at most 2 vertices that overlap with u in k positions. Thus,

$$\begin{aligned} P[\bigcup_{(u,v)\in J_k, k>0} B_{uv}] &\leq \sum_{(u,v)\in J_{2\log n+1}} P[B_{uv}] + \sum_{(u,v)\in J_{2\log n+2}} P[B_{uv}] \\ &\leq 2n(\frac{1}{4n}) = \frac{1}{2}. \end{aligned}$$

Therefore, $P[\bigcup_{(u,v)\in VxV} B_{uv}] \leq P[\bigcup_{(u,v)\in J_0} B_{uv}] + P[\bigcup_{(u,v)\in J_k, k>0} B_{uv}] \leq \frac{9}{16}$ and our result follows. This implies the existence of a 2-labeling of C_n that is an $\mathrm{LD}^{\lceil \log n+1 \rceil}$ labeling. $\square$

5 LD^1 for Almost All Graphs, Open Problems, Future Directions

This paper has explored the local distinguishing number of cycles. In particular, we showed that $\mathrm{LD}^i(C_n)$ is monotonic when $i = 1$. However, our general upper bound on $\mathrm{LD}^i(C_n)$ does not. Very recently, Alon [2], has improved the upper bound on $\mathrm{LD}^i(C_n)$ to within a factor of 2 by using de Bruijn sequences and a result of Lempel [5]. Again, this still does not resolve monotonicity. We make the following conjecture.

Conjecture 5.1. *For a constant $i > 1$, $LD^i(C_n)$ is monotonically non-decreasing in n.*

We are also interested in the local distinguishing number of other graph families, and for general graphs. We, thus, ask the following open question:

Question 5.2. *What is the ith local distinguishing number of H_n, the n-dimensional hypercube?*

We remark that a random graph property Babai, Erdös and Selkow used for testing random graph isomorphism in [3] gives a good starting point for the investigation of $\mathrm{LD}^1(G)$ for general graphs.

Consider a graph, H, on n vertices where all but two of the vertices have distinct degrees. $\mathrm{LD}^1(H) \leq 2$ since only the two vertices with the same degrees need to be distinguished. For K_n, $\mathrm{LD}^1(K_n) = n$. While we do not have a tight upper bound for $\mathrm{LD}^1(G)$ for an arbitrary graph G, we show that for almost all graphs, $\mathrm{LD}^1(G) = O(\log n)$.

Let G be a random graph on n vertices where the pair (u, v) is chosen independently as an edge of G with probability p. Label the vertices of G as $v_1, v_2, ..., v_n$ such that $d(v_i) \geq d(v_{i+1})$. Let $r = \lfloor 3 \log_2 n \rfloor$. Consider the following conditions:

(i) Let $U = \{v_1, v_2, ..., v_r\}$. All vertices in U have distinct degrees;

(ii) The remaining $n - r$ vertices in $V - U$ have distinct neighborhoods in U.

Babai, et al. showed that for sufficiently large n, with probability greater than $1 - n^{-\frac{1}{7}}$ a random graph on n vertices satisfy the two conditions above.

Theorem 5.3. *Almost all graphs can be* 1*-locally distinguished with* $\lfloor 3 \log n + 1 \rfloor$ *colors.*

Proof: Let G be a random graph that satisfies conditions (i) and (ii). To distinguish vertices in U from vertices in $V - U$, color all vertices in $V - U$ with color 1 and the vertices in U with colors $2, ..., \lfloor 3 \log n \rfloor + 1$. The vertices in $V - U$ can be distinguished from each other based on their neighborhood sets in U, while the vertices in U can be distinguished from each other based on their vertex degrees. $\square$

The tightness of Theorem 5.3 is unknown and we leave it as an open question.

6 Acknowledgements

We thank Mike Albertson for his encouragement and an early draft of [1]. We also thank the anonymous referees for helpful comments that improved the presentation of this paper. CTC and LJC were supported in part by ONR grant N00014-96-1-0829.

References

[1] M. Albertson, K. Collins, Symmetry Breaking in Graphs, *Electron. J. of Combin.*, 3 (1996).

[2] N. Alon, personal communication.

[3] L. Babai, P. Erdös, S. Selkow, Random Graph Isomorphism, *SIAM J. Computing*, 3 (1980) 628–635.

[4] B. Bollobás, Degree sequences of Random Graphs, *Trans. Amer. Math. Soc.*, 267 (1981) 41–52.

[5] A. Lempel, m-ary closed sequences, *J. Comb. Theory Ser. A* 10 (1971), 253–258.

[6] R. Motwani, P. Raghavan, *Randomized Algorithms*, Cambridge University Press (1995).

[7] D. West, *Introduction to Graph Theory*, Prentice–Hall (1996).