Randomized versus Nondeterministic Communication Complexity *

Paul Beame
Joan Lawry
Department of Computer Science and Engineering, FR-35
University of Washington
Seattle, Washington 98195

Abstract

Our main result is the demonstration of a Boolean function f with nondeterministic and co-nondeterministic complexities $O(\log n)$ and ϵ-error randomized complexity $\Omega(\log^2 n)$, for $0 \leq \epsilon < 1/2$. This is the first separation of this kind for a decision problem.

1 Introduction

The two-party communication complexity of Boolean functions has been studied extensively since [Yao79]. The ground-breaking work of [KW88] connecting the depth complexity of Boolean functions with the two-party communication complexity of related search problems has sparked renewed interest in the subject. Furthermore, [Raz88] and [RW90] have shown that lower bounds for these related search problems may sometimes be found by reduction from known lower bounds in the standard model of two-party computation of decision problems.

Many variants of the standard model have been analyzed. For notation, let $N_1(f)$ be the nondeterministic complexity of a function f, $N_0(f)$ be its co-nondeterministic complexity—i.e., $N_0(f) = N_1(\bar{f})$, $N(f)$ be the maximum of $N_1(f)$ and $N_0(f)$, $D(f)$ be its deterministic complexity, and $R_\epsilon(f)$ be its ϵ-error randomized complexity. $R_\epsilon(f)$ is notable because of its use in [RW90] to obtain very strong lower bounds on circuit depth.

The relationships among these complexities are interesting as well. Clearly, $D(f) \geq R_0(f) \geq N(f)$, and it is easy to find Boolean functions f for which $R_\epsilon(f)$ is substantially below $N(f)$ [Yao81]. A somewhat surprising result is that for all Boolean functions $D(f) = O(N_1(f)N_0(f))$ [AUY83], which was improved to $D(f) \leq N_1(f)N_0(f)(1 + o(1))$ by [HR88]. [Für87], [HR88], [Raz88] also demonstrated functions for which this bound is tight, which bettered a separation shown in [MS82]. [Für87] actually improved the randomized upper bound of [MS82] to show that there exists a Boolean function with $R_0(f) = O(N(f))$ and $D(f) = \Omega(R_0^2(f))$. Our primary question is whether there exists a Boolean function with 0-error randomized complexity bounded away from its nondeterministic complexity. Of particular interest is the range of $R_0(f)$ between $D(f)$ and $N(f)$ when $D(f) = \Omega(N_1(f) \cdot N_0(f))$. Is there a function with $R_0(f)$ bounded as far away from $N(f)$ as its deterministic complexity? What about ϵ-error randomized complexity? Despite considerable study of communication complexity these questions have not been addressed previously.

Our main result is that there exists a Boolean function f with $N_1(f) = N_0(f) = O(\log n)$ and $R_\epsilon(f) = \Omega(\log^2 n)$, for $0 \leq \epsilon < 1/2$. Our decision problem is based on the following two-player n-node graph game: the path player has an st-path of length, e.g. $n^{1/150}$, and the color player has a 01-coloring of the n nodes, where together the path-coloring input pair is such that the path has exactly one bi-chromatic edge. The players cooperatively try to determine the parity of the bi-chromatic edge, where edges are numbered according to their position in the input path.

*Research supported by the National Science Foundation, under grants CCR-8858799 and CCR-8907960

24th ANNUAL ACM STOC - 5/92/VICTORIA, B.C., CANADA
© 1992 ACM 0-89791-512-7/92/0004/0188...$1.50

The obvious nondeterministic solution has complexity $2\log n + 2$, for either a 1 or a 0 solution: the color player guesses which edge of the path is the bi-chromatic path edge, sends the edge number and the associated pair of vertices to the path player, who then verifies that the edge vertices have different colors and outputs the parity of the edge number. It is not too hard to show that these are asymptotically the best possible bounds for $N_1(f)$ and $N_0(f)$.

The deterministic complexity of this problem is $\Theta(\log^2 n)$. For the upper bound, the path player simply binary searches his path for the bi-chromatic edge, repeatedly querying the color player for the color of a given node. The lower bound follows from [KW88].

To bound randomized complexity away from nondeterministic complexity, the static methods previously used for probabilistic lower bounds on decision problems are insufficient since they can at best show a lower bound of $N(f)$ on the randomized complexity of f: The rectangle approach of [Yao79] as extended in [Yao83] for randomized complexity requires one to show that, under some distribution on the inputs, either every 0- or 1-rectangle with sufficiently high probability cannot have a small probability of error. However, if $N(f)$ is small then there exists an entirely correct cover of all possible inputs that contains very large 0- and 1-rectangles.

Instead, we base our result on the iterative restriction techniques used to obtain deterministic lower bounds for search problems in [KW88] and [GH89], as well as on the extension of the latter by [RW89] to an ϵ-error randomized lower bound. The general iterative approach consists of maintaining sets of candidate inputs for each player, where each input is consistent with the communication so far and the input sets have a "nice" structure. In addition, for probabilistic arguments, the structure must be such that, for many rounds of communication, the input sets retain a reasonable fraction of inputs for which the protocol answers correctly. For search problems this reasonable fraction is quite low, since the many possible choices for the answer are likely to prevent a correct guess. For a decision problem, however, this fraction of correct answers must be more than one-half. Furthermore, care must be taken to maintain a roughly even balance of the two possible answers.

The questions we address have been previously considered in models where a similar quadratic upper bound applies: For all Boolean functions, [BI87], [HH87], [Tar88] established $D^d(f) \leq N_1^d(f)N_0^d(f)$ in the decision tree model. [DF89] noted the similarity between these quadratic upper bounds and developed a multi-party communication game that is a generalization of both the 2-party game and decision trees. They were only able to obtain a somewhat weaker upper bound on deterministic multi-party complexity of the same type, but [Law92] has shown that a roughly quadratic upper bound does apply in this model to the 0-error randomized complexity. We can generalize our questions above to whether or not this latter bound is tight. We do have some evidence for this in the decision tree model. [SW86] show that a function given by [Sni85] has 0-error randomized complexity bounded away from $N^d(f)$: $N^d(f) = D^d(f)^{.5}$ and $R_0^d(f) = \Theta(D^d(f)^{.753\ldots})$. It is also interesting to observe that for search problems the gaps between nondeterministic, randomized, and deterministic complexities can be arbitrarily large under both the two-party communication model [RW90] and the decision tree model [LNNW91].

We should finally note that we have not quite shown that the bounds of [AUY83] and [HR88] on $R_0(f)$ are the best possible. Our ϵ-error randomized lower bound applies to inputs that are cross products, but the nondeterministic upper bounds may not hold for either $N_1(f)$ or $N_0(f)$, depending on how the problem is extended to a cross product. While the deterministic upper bound of [AUY83], [HR88] doesn't apply to non-cross products, we do, however, have the quadratic relationship between deterministic and nondeterministic complexities for our problem: $D(f) = \Theta(N_1(f)N_0(f))$. Thus our result bounds $R_\epsilon(f)$ as far away from $N(f)$ as possible—at $D(f)$.

2 Overview

Given any protocol for the path-coloring problem our goal is to prove the existence of an input pair that forces the players to communicate many bits

before they can determine the solution. We follow [Yao81] and work with a deterministic protocol that errs on some fixed fraction of the input. The *quality* of the protocol on a given input domain is the fraction of the input for which the protocol answers correctly. For our decision problem, we must ensure that the quality is bounded above one-half.

We base our lower bound strategy on that of [RW89]. The proof of [RW89] extended the deterministic result of [KW88] to an ϵ-error randomized lower bound by generalizing to a well behaved, non-cross product input and accounting for input pairs on which the protocol errs.

The strategy maintains a set of candidate inputs for each player that are consistent with the communication sent so far. In addition, the strategy eliminates some of the eligible inputs in the interest of imposing a favorable structure on the input sets—one that precludes an easy solution. The relevant quantities to preserve are set density and quality. It turns out that the set density of the colorings is very easy to maintain at a sufficiently high level, and it is primarily the set density of the paths that is at issue.

In order to admit a nontrivial lower bound, the strategy must compensate for some of the effect of the communication. Since set density is the quantity most adversely affected, the main objective of the strategy is to restore path set density, while maintaining sufficiently high quality. The primary tool for increasing set density is reduction of each input path to a relatively large subpath—either the left or the right side of the path. After this reduction, however, a clean-up of the input sets is necessary to ensure that each reduced pair is from a valid, high-quality pair from the original input—and that the solution edge lies within the reduced input.

One of the necessary conditions for improving set density by path reduction is a large collection of paths with high quality on both sides—that is, both in the restricted path and its complement. If we have this large collection of paths, a single path restriction results in a large gain in set density: the path is reduced to the side of high quality for many paths and the clean-up step ensures valid input pairs. On the other hand, if we don't have enough paths with high quality on both sides, then we show that applying path restriction in a different way improves average quality and, by doing this repeatedly, overall quality eventually becomes high enough that the density-increasing restriction can be applied. Paradoxically, each step of path restriction for this purpose incurs a loss in set density, which the strategy must also try to offset with the density-increasing restriction. A clean-up also follows each of these quality-increasing reductions, though a couple of additional steps of refinement to the input are necessary before the clean-up process can be applied.

One consequence of our work is a significant simplification of the proof in [RW89]. Furthermore, our analysis differs from the simplified version of their proof in that our lower bound strategy cannot tolerate the low quality levels that are adequate for the search problem. We compensate for this difference by giving up subpath length in our restrictions. As a result, we argue termination of the iterative quality-increasing path restrictions based on quality alone, as opposed to the product of density and quality in [RW89], and we argue about path length itself as opposed to the product of the length and the 100th power of the quality. This further streamlines some of the argument and, additionally, affords application of the amortization technique of [GH89] to bound parameters which rely on the number of repetitions in a phase of quality-increasing restrictions.

3 Definitions

We use the following decision problem in the two-party communication model: Let $\mathcal{D}$ be a domain of n vertices over which two sets P_n^l and C_n are defined, where P_n^l is the set of all simple paths from vertex s to vertex t of length l, $n^{1/200} \leq l < (1/2)n^{1/100}$, and C_n is the set of all 01-colorings on the n vertices. Let $P \subseteq P_n^l$ be the set of candidate inputs for the path player and $C \subseteq C_n$ be the set of candidate inputs for the color player, where P and C are determined by the communication of the protocol and by restrictions of the lower bound strategy. Additionally, any actual path-coloring input pair has exactly one bi-chromatic edge. Then, the value of the function is the parity of the bi-

chromatic edge, where edges are numbered according to their position in the input path.

More formally, the input domain is defined as $\Phi(P,C) = \{(p,c) \in P \times C$: exactly one edge of p is bi-colored by $c\}$, where $\varphi(P,C) = |\Phi(P,C)|$. The set density of the paths and the colorings are defined as $\mu(P) = |P|/|P_n^l|$, $\mu(C) = |C|/|C_n|$. The protocol is guaranteed to behave correctly on only some fraction of the valid input pairs. We denote this fraction by $\gamma(P,C)$.

Since the input domain is not $P \times C$, the path splitting process will necessitate considering both $\varphi(P,C)$ and $\gamma(P,C)$ as weighted sums. To facilitate this, let $\Phi^i(P,C) = \{(p,c) \in \Phi(P,C)$: edge i of p is bi-colored by $c\}$, with $\varphi^i(P,C) = |\Phi^i(P,C)|$. Define $\gamma^i(P,C)$ to be the fraction of input pairs in $\Phi^i(P,C)$ on which the protocol is correct.

Let $\mathcal{L} = \{1,\ldots,l_L\}$, $\mathcal{R} = \{l_L+1,\ldots,l\}$, $l_L + l_R = l$, be a bi-partition of the path length l into a left and a right side. Define $\gamma_L(P,C)$ and $\gamma_R(P,C)$ to be the average quality of the protocol on inputs whose answer falls in the left and the right sides, respectively, of paths in P. That is, for $S \subseteq \{1,\ldots,l\}$,

$$\gamma_S(P,C) = \frac{\sum_{i\in S} \gamma^i(P,C)\varphi^i(P,C)}{\sum_{i\in S} \varphi^i(P,C)}.$$

Also, define $\gamma_S(p,C)$ to be $\gamma_S(\{p\},C)$.

Throughout the strategy, for the current values of l_L, l_R, s, t, and n, let X be the set of all paths of length l_L on n vertices with first vertex s and Y be the set of all paths of length l_R on n vertices with last vertex t. Furthermore, let $p = (x;y)$ denote a path consisting of a left subpath $x \in X$ appended to a right subpath $y \in Y$. For each $x \in X$, let $\delta(x) = |\{y \in Y : (x;y) \in P\}|/|Y|$ be the fraction of valid extensions of x that form a path in P. Similarly, for each $y \in Y$, let $\delta'(y) = |\{x \in X : (x;y) \in P\}|/|X|$ be the fraction of valid extensions of y that form a path in P.

4 The Strategy

From the point of view of the lower bound strategy, a round consists of β bits sent between the two players, in any alternation. Later, we will set the number of bits that comprise a round to optimize the lower bound, subject to constraints that arise on β. The strategy of the lower bound adversary consists of the following steps for each round:

1. **Select β bits to be sent during the round such that these bits correspond to a subset of the candidate inputs that retains relatively high set density and quality.**

   $(P,C) \leftarrow$ some $(P' \subseteq P, C' \subseteq C)$ such that $\gamma(P',C') \geq (1-2^{-\beta})\gamma(P,C)$, $\mu(P') \geq 2^{-3\beta}\mu(P)$ and $\mu(C') \geq 2^{-3\beta}\mu(C)$, where (P',C') is consistent with the bits chosen.

2. **Partition the path into the subpath that will become the restricted path and the complement of the subpath.**

   Determine a bi-partition $\mathcal{L} = \{1,\ldots,l_L\}$, $\mathcal{R} = \{l_L+1,\ldots,l\}$, $l_L + l_R = l$, of the path length l, such that $l_L, l_R \geq \frac{l}{4k}$, and $\gamma_L(P,C), \gamma_R(P,C) \geq (1-\frac{1}{k})\gamma(P,C)$.

3. **For each side of the partition, form the subset of the input set consisting of paths on which the protocol performs well when the bi-colored edge lies on the side in question.**

   $$P_L \leftarrow \{p \in P : \gamma_L(p,C) \geq (1-\tfrac{8}{k})\gamma_L(P,C)\},$$
   $$P_R \leftarrow \{p \in P : \gamma_R(p,C) \geq (1-\tfrac{8}{k})\gamma_R(P,C)\}.$$

4. **While less than $1/2$ of the paths have high quality on both sides, iteratively and selectively restrict the path length to increase average quality.**

   The density of paths having high quality on one side is too low to apply the density-increasing reduction to either side. We repeatedly increase $\gamma(P,C)$ by restricting the path set to members with high quality on the side of the low density set. Eventually, the average quality will be high enough to guarantee that many paths will have high quality on both sides.

   One of the sets is small; either $|P_L| < \frac{3}{4}|P|$ or $|P_R| < \frac{3}{4}|P|$, say $|P_L|$.

   (a) **Keep only those paths whose contribution to $\varphi_L(P,C)$ is about average.**

This allows accounting for the non-cross product nature of the input in subsequent calculations. Let $\phi = \frac{\varphi_L(P,C)}{|P|}$.

$$P \leftarrow \{p \in P : \varphi_L(p,C) = \phi(1 \pm O(n^{-\frac{1}{10}}))\}.$$

(b) **Restrict the input to high quality paths.**

$$P \leftarrow \{p \in P : \gamma_L(p,C) \geq (1 + \tfrac{2}{k})\gamma_L(P,C)\}.$$

Let $\Gamma = (1 + \frac{1}{2k})\gamma_L(P,C)$ denote the high minimum per-path left quality achieved in this step. We will call the left subpaths *stems* and the right subpaths *extensions*. We say a stem has many extensions if the stem is a left subpath of many paths in P. Eventually, we will restrict the set of paths to a subset of the stems.

(c) **Eliminate paths whose stem is not the stem of many paths.**

This step refines the input set for condition 1 of Lemma 4, enabling Step 4d.

$$P \leftarrow \{p = (x;y) \in P : \delta(x) \geq n^{-\frac{1}{100}}\}.$$

Denote the set of stems by

$$X^P \leftarrow \{x \in X : \exists p = (x;y) \in P\}.$$

(d) **Clean-up and restrict input sets.**

$\mathcal{D} \leftarrow \mathcal{D}' \subset \mathcal{D}$, where $\mathcal{D}'$ is a new domain set of $n' = n - \sqrt{n}$ vertices so that the following may be found:

$(P,C) \leftarrow$ some $(P' \subseteq P^{l_L}_{n'}, C' \subseteq C_{n'})$ over domain $\mathcal{D}'$ such that

- $\mu'(P') = |P'|/|P^{l_L}_{n'}| \geq |X^P|/(10|X|)$,
- $\mu'(C') = |C'|/|C_{n'}| \geq (1 - O(n^{-1/10}))\mu(C)$,
- and $\gamma_L(P',C') \geq (1 - O(n^{-1/10}))\Gamma$;

$l \leftarrow l_L$; $n \leftarrow n'$.

(e) **Partition the path and, for each side of the partition, form the subsets of high-quality paths.**

That is, execute Steps 2 and 3; repeat Step 4, if necessary.

5. **Once 1/2 of the paths have high quality on both halves, restrict the input to a subset of the intersection of the two filtered sets, thereby increasing path set density.**

The strategy is now in a position to substantially increase the density of the set of remaining paths, which will be accomplished by restricting the path length. First, however, we keep only those paths with high quality on both sides.

$$P \leftarrow P_L \bigcap P_R.$$

Now, the strategy decides which subpath will become the reduced path. Consider the left and right sides of paths in P. Let

$$X^P \leftarrow \{x \in X : \delta(x) \geq \mu(P)/4\},$$

$$Y^P \leftarrow \{y \in Y : \delta'(y) \geq \mu(P)/4\}.$$

Either $|X^P|/|X| \geq \sqrt{\mu(P)/2}$ or $|Y^P|/|Y| \geq \sqrt{\mu(P)/2}$. Say $|X^P|/|X| \geq \sqrt{\mu(P)/2}$.

$$P \leftarrow \{p = (x;y) \in P : x \in X^P\}.$$

6. **Keep only those paths whose contribution to $\varphi_L(P,C)$ is about average.**

Same as Step 4a.

7. **Clean-up and restrict path length.**

Same as Step 4d. Note that here, the minimum per-path quality is $\gamma_L = (1 - \frac{9}{k})\gamma(P,C)$, which was determined by the most recent execution of Steps 2 and 3.

5 Lemmas for Structured, Non-Cross Product Inputs

For the analysis of the lower bound strategy we need to quantify the effects of each step on $\mu(P)$, $\mu(C)$, and $\gamma(P,C)$. We would like to consider P and C in isolation, but to do so we would need to assume that the input is a cross-product of the two sets. However, even though our input is not a cross product, for our purposes the input behaves as if it were, according to the following results from [RW89]:

Lemma 1 Let $l < n^{1/100}, \mu(P) \geq n^{-1/100}$, and $\mu(C) \geq 2^{-n^{1/100}}$. Then, for all i, $\varphi^i(P,C) = \mu(P)\mu(C)\varphi^i(P_n^l, C_n)(1 \pm O(n^{-1/10}))$.

Corollary 2 Let $l < n^{1/100}, \mu(P) \geq n^{-1/100}$, and $\mu(C) \geq 2^{-n^{1/100}}$. Then, for all $1 \leq i,j \leq l$, $\varphi^i(P,C) = \varphi^j(P,C)(1 \pm O(n^{-1/10}))$.

Corollary 3 Let $l < n^{1/100}, \mu(P) \geq n^{-1/100}$, and $\mu(C) \geq 2^{-n^{1/100}}$. Then $\varphi(P,C) = \mu(P)\mu(C)\varphi(P_n^l, C_n)(1 \pm O(n^{-1/10}))$.

We also use the following lemma from [RW89] to "clean-up" the input sets after reducing the path set to a subset of what it was previously. This result ensures that the shortened input paths each have a valid, high quality extension in the original set of paths that is consistent with the current restricted set of colorings.

Lemma 4 Let $\mathcal{D}$ be a domain of n vertices over which the sets of paths P_n^l, where $l < n^{-1/100}$, and the set of colorings C_n are defined. Let $P \subseteq P_n^l$ and $C \subseteq C_n$, such that for all $p = (x;y) \in P$, where the length of x is l_L, $l_R = l - l_L$,

$$\delta(x) \geq n^{-1/100}, \quad (1)$$

$$\frac{\gamma_L(p,C)\varphi_L(p,C)}{|C_{n-l_R}|} > 2^{-n^{1/100}}. \quad (2)$$

Then, there exists $\mathcal{D}' \subset \mathcal{D}$, where $\mathcal{D}'$ is a set of $n' = n - \sqrt{n}$ vertices, $P' \subseteq P$, $C' \subseteq C$, and $\lambda = (1 - O(n^{-\frac{1}{10}}))$ such that

- all $c \in C'$ agree on the colors of $\mathcal{D} \setminus \mathcal{D}'$,
- for all $x \in X$ such that there is some $p = (x;y) \in P'$ this y is unique, the vertices of x are in $\mathcal{D}'$, the vertices of y are in $\mathcal{D} \setminus \mathcal{D}'$ and colored the same as t for all $c \in C'$, and

$$\gamma_L(p,C')\varphi_L(p,C') \geq \gamma_L(p,C)\varphi_L(p,C)\frac{|C_{n'}|\lambda}{|C_{n-l_R}|},$$

- $\mu'(P') = |P'|/|P_{n'}^{l_L}| \geq |X^P|/(5|X|)$,
- letting $\mu'(C') = |C'|/|C_{n'}|$ we have $\mu(C)/\lambda \geq \mu'(C') \geq \mu(C)\lambda$.

Proof [RW89], Lemma 2.4, Lemma 2.3, and a slightly modified version of Claim 3.2. □

6 The Lower Bound

We now bound the changes in $\mu(P)$, $\mu(C)$, $\gamma(P,C)$, and l for the steps of the lower bound strategy. In what follows, we give full details for each step of the strategy for an $\Omega(\log^2 n/\log\log n)$ lower bound. We also sketch the modifications that give the $\Omega(\log^2 n)$ bound.

The following lemma describes how in Step 1 the lower bound strategy limits the loss in set density and quality due to the communication.

Lemma 5 Suppose that, at the beginning of Step 1, the path player has some input from P and the color player has some input from C, such that $\mu(P) \geq 2^{3\beta}n^{-1/100}$, $\mu(C) \geq 2^{3\beta-n^{1/100}}$, and $\gamma(P,C) > 1/2$. Let the path and the cut player together send at most $\beta > 1$ bits in the round. Then, there exist $P' \subseteq P$ and $C' \subseteq C$ consistent with the communication such that $\mu(P') \geq 2^{-3\beta}\mu(P), \mu(C') \geq 2^{-3\beta}\mu(C)$, and $\gamma(P',C') \geq (1 - 2^{-\beta})\gamma(P,C)$.

Proof Let α be the fraction of path-coloring pairs (p,c) that are in classes (P',C') such that $\gamma(P',C') \geq (1 - 2^{-\beta})\gamma(P,C)$. α is smallest when those classes with quality at least $(1 - 2^{-\beta})\gamma(P,C)$ actually have perfect quality and the remaining $1 - \alpha$ fraction all have quality just less than $(1 - 2^{-\beta})\gamma(P,C)$. From this, $\gamma(P,C) \leq \alpha + (1-\alpha)(1 - 2^{-\beta})\gamma(P,C)$, which implies $\gamma(P,C)/2^{\beta} \leq \alpha$. Since this fraction is non-zero, there exists some class (P',C') such that $\gamma(P',C') \geq (1 - 2^{-\beta})\gamma(P,C)$.

We show that one of these classes (P',C') is large enough by partitioning α into an α_s fraction of pairs in classes (P',C') with $\mu(P')\mu(C') < 2^{-3\beta}\mu(P)\mu(C)$ and an α_l fraction of pairs in classes (P',C') with $\mu(P')\mu(C') \geq 2^{-3\beta}\mu(P)\mu(C)$. It suffices to show that α_l is nonzero, and we do so by showing $\alpha_s < \gamma(P,C)/2^{\beta}$.

Consider some class (P',C') with $\mu(P')\mu(C') \leq 2^{-3\beta}\mu(P)\mu(C)$. If $\mu(P') \geq n^{-1/100}$ and $\mu(C') \geq 2^{-n^{1/100}}$, we can apply Lemma 1 to show that

$$\varphi(P',C')/\varphi(P,C) \leq \lambda\mu(P')\mu(C')/(\mu(P)\mu(C)),$$

where $\lambda \leq 1 + cn^{-1/10}$, for some constant c. Thus the contribution of any such class to α_s cannot be more than a $2^{-3\beta}\lambda$ fraction of all pairs in $\Phi(P,C)$.

If $\mu(P') < n^{-1/100}$ then we can apply Lemma 1 to show that

$$\begin{aligned}\varphi(P',C') &\leq n^{-1/100}\mu(C)\varphi(P_n^l, C_n)\lambda \\ &\leq (n^{-1/100}/\mu(P))\varphi(P,C)\lambda',\end{aligned}$$

where $\lambda' \leq 1 + c'n^{-1/10}$, for some constant c'. By the conditions on $\mu(P)$, $n^{-1/100}/\mu(P) \leq 2^{-3\beta}$, and any such class contributes at most a $2^{-3\beta}\lambda'$ fraction of $\Phi(P,C)$. If $\mu(C') < 2^{-n^{1/100}}$ then similar reasoning shows that $\varphi(P',C') \leq 2^{-3\beta}\varphi(P,C)\lambda$. In each case any such class contributes at most a $2^{-3\beta}(1 + O(n^{-\frac{1}{10}}))$ fraction of pairs in $\Phi(P,C)$. Since there are at most 2^{β} such classes,

$$\alpha_s \leq 2^{-2\beta}(1 + O(n^{-\frac{1}{10}})) < 2^{-\beta-1} < \gamma(P,C)/2^{\beta},$$

as required. □

The following lemma ensures that the strategy is always able to execute Step 2; that is, it guarantees the existence an appropriate bi-partition of the path length.

Lemma 6 Let $\mu(P) \geq n^{-1/100}$, and $\gamma(P,C) > 1/2$ be the path set density and protocol quality on the left side just prior to step 2 of the lower bound strategy. In addition, let the paths in P be of length $l \geq 4k$ and $\mu(C) \geq 2^{-n^{1/100}}$ be the coloring set density. Then, there exist l_L and l_R, $l_L + l_R = l$, such that

(a) $l_L, l_R \geq l/(4k)$ and $\gamma_L(P,C), \gamma_R(P,C) \geq (1 - 1/k)\gamma(P,C)$, and

(b) if the smaller of l_L and l_R, say l_R, is at most l/m then $\gamma_R(P,C) \geq (1 + (m-3)/3k)\gamma(P,C)$.

Proof Part (a): From the definitions of $\gamma(P,C)$, $\gamma_L(P,C)$, $\gamma_R(P,C)$, and Corollary 2, we know:

$$\begin{aligned}l\gamma(P,C) &= \sum_{i \in \mathcal{L}\cup\mathcal{R}} \gamma^i(P,C)(1 \pm O(n^{-\frac{1}{10}})), \\ l_L\gamma_L(P,C) &= \sum_{i \in \mathcal{L}} \gamma^i(P,C)(1 \pm O(n^{-\frac{1}{10}})), \\ l_R\gamma_R(P,C) &= \sum_{i \in \mathcal{R}} \gamma^i(P,C)(1 \pm O(n^{-\frac{1}{10}})).\end{aligned}$$

Therefore,

$$\begin{aligned}l_L\gamma_L(P,C) + l_R\gamma_R(P,C) &= l\gamma(P,C)(1 \pm O(n^{-\frac{1}{10}})) \\ &\geq l\gamma(P,C)\lambda,\end{aligned}$$

where $\lambda \geq 1 - cn^{-1/10}$, for some constant c.

First try $l_L = \lfloor l/2 \rfloor$. If both $\gamma_L(P,C)$ and $\gamma_R(P,C)$ are at least $(1 - 1/k)\gamma(P,C)$ then we are done. Otherwise one of them is smaller, say $\gamma_L(P,C) < (1 - 1/k)\gamma(P,C)$. Now consider the smallest $l_L > l/2$ such that the average quality $\gamma_L(P,C) \geq (1 - 1/k)\gamma(P,C)$. We will show that $l_R \geq l/4k$ and $\gamma_R(P,C) \geq (1 - 1/k)\gamma(P,C)$.

From our construction of l_L and l_R,

$$l_L(1-1/k)\gamma(P,C) + 1 + (l - l_L)\gamma_R(P,C) \geq l\gamma(P,C)\lambda,$$

where the $+1$ term compensates for the discreteness of l_L. Since $l \geq 4k$ and $\gamma(P,C) > 1/2$ we have $1 < l\gamma(P,C)/2k$. Using this substitution and collecting like terms in the above inequality we have

$$\begin{aligned}&(l - l_L)\gamma_R(P,C) \\ &\quad\geq [(\lambda - 1/2k)l - (1 - 1/k)l_L]\gamma(P,C) \\ &\quad= (1 - 1/k)(l - l_L)\gamma(P,C) \\ &\qquad\quad +(1/2k + \lambda - 1)l\gamma(P,C) \\ &\quad\geq (1 - 1/k)(l - l_L)\gamma(P,C) + (l/3k)\gamma(P,C).\end{aligned}$$

Thus $\gamma_R(P,C) > (1 - 1/k)\gamma(P,C)$.

Using $\gamma_R(P,C) \leq 1$, $l_R = l - l_L$, and $\gamma(P,C) \geq 1/2$,

$$\begin{aligned}l_R &\geq (1 - 1/k)l_R\gamma(P,C) + (l/3k)\gamma(P,C) \\ &> (1 - 1/k)l_R/2 + l/6k.\end{aligned}$$

Therefore $l < 3k(1 + 1/k)l_R < 4kl_R$, and so $l_R \geq l/4k$ as required. □

Step 4a is necessary for relating path qualities weighted by the contribution of the paths to $\varphi_L(P,C)$ to the average quality of the path set.

Lemma 7 Let $\mu(P) \geq 4n^{-1/100}$ be the path set density, $\mu(C) \geq 2^{-n^{1/100}}$ be the coloring set density, and $\gamma_L(P,C) > 1/2$ be the protocol quality on the left side just prior to Step 4a of the lower bound strategy, and $\mu(P')$, $\gamma_L(P',C)$ be the subsequent density and quality. Then, $\mu(P') \geq \mu(P)/2$ and $\gamma_L(P',C) \geq \gamma_L(P,C)(1 - O(n^{-\frac{1}{10}}))$.

Proof More specifically, Step 4a does the following: $P \leftarrow \{p \in P : \phi(1 - 2cn^{-\frac{1}{10}}) \leq \varphi_L(p,C) \leq \phi(1 + 2cn^{-\frac{1}{10}})\}$, where c is the constant of Lemma 1

and $\phi = \varphi_L(P,C)/|P|$. Alternately, ϕ can be expressed as $\phi = l_L\mu(C)|C_n|/2^l$, since ϕ is the average number of colorings per path that color one edge in l_L with both 0 and 1 and color all of l_R the same color as t. We show that the fraction of paths eliminated by this filter is small.

Let $P_< = \{p \in P : \varphi_L(p,C) < \phi(1 - 2cn^{-\frac{1}{10}})\}$. By definition, for all paths $p \in P_n^l$, $\varphi_L(p,C_n) = l_L|C_n|/2^l$. Therefore, $\varphi_L(P_n^l,C_n) = l_L|P_n^l||C_n|/2^l$. Suppose $\mu(P_<) \geq n^{-1/100}$. Then, by applying Lemma 1,

$$\begin{aligned}
&\varphi_L(P_<,C) \\
&\quad \geq \mu(P_<)\mu(C)\varphi_L(P_n^l,C_n)(1 - cn^{-\frac{1}{10}}) \\
&\quad \geq \mu(P_<)\mu(C)\frac{l_L|P_n^l||C_n|}{2^l}(1 - cn^{-\frac{1}{10}}).
\end{aligned}$$

However, according to the definition of $P_>$,

$$\begin{aligned}
\varphi_L(P_<,C) &< |P_<|\phi(1 - 2cn^{-\frac{1}{10}}) \\
&< \mu(P_<)|P_n^l|\frac{l_L\mu(C)|C_n|}{2^l}(1 - 2cn^{-\frac{1}{10}}),
\end{aligned}$$

which is a contradiction.

Similarly, we can argue that $P_> = \{p \in P : \varphi_L(p,C) > \phi(1 + 2cn^{-\frac{1}{10}})\}$ must have $\mu(P_>) < n^{-1/100}$. Thus $\mu(P_> \cup P_<) < 2n^{-1/100}$, and $\mu(P') \geq \mu(P) - 2n^{-1/100} \geq \mu(P)/2$.

Since either $\mu(P')$ or $\mu(P \backslash P') \geq n^{-1/100}$, where $P \backslash P' = P_\leq \cup P_\geq$, we can apply Lemma 1 to derive the following:

$$\begin{aligned}
&\gamma_L(P,C)\varphi_L(P,C)(1 - O(n^{-\frac{1}{10}})) \\
&\quad \leq \gamma_L(P',C)(1 - \frac{\mu(P\backslash P')}{\mu(P)})\varphi_L(P,C) \\
&\qquad + \gamma_L(P\backslash P',C)\frac{\mu(P\backslash P')}{\mu(P)}\varphi_L(P,C).
\end{aligned}$$

$\gamma_L(P',C)$ is lowest when the eliminated paths have perfect quality. Hence,

$$\begin{aligned}
&\gamma_L(P,C)(1 - O(n^{-\frac{1}{10}})) \\
&\quad \leq \gamma_L(P',C)(1 - \frac{\mu(P\backslash P')}{\mu(P)}) + \frac{\mu(P\backslash P')}{\mu(P)},
\end{aligned}$$

from which we get

$$\begin{aligned}
\gamma_L(P',C) &\geq \gamma_L(P,C)(1 - O(n^{-\frac{1}{10}})) - 2n^{-1/100} \\
&= \gamma_L(P,C)(1 - O(n^{-\frac{1}{10}})),
\end{aligned}$$

since $\gamma_L(P,C) \geq 1/2$. □

The next lemma precisely quantifies the increase in quality—and decrease in set density—due to Step 4b's restriction of the input to the high quality paths in the low density set formed in Step 3. Step 4b is the only way the lower bound strategy increases average quality, and the increase is enough that Step 4 has a net gain in quality. Thus, as γ increases, more and more paths have high quality on both sides, and eventually the path set will meet the termination conditions for Step 4.

Lemma 8 Let $\mu(P) \geq n^{-1/100}$, $\mu(C) \geq 2^{-n^{1/100}}$, and $\gamma_L(P,C) > 1/2$ be the path set density, coloring set density, and protocol quality on the left side just prior to step 4b of the lower bound strategy, and let $\mu(P')$ and $\gamma_L(p',C)$ be the subsequent path set density and per-path left protocol quality. Then,

(a) $\mu(P') \geq \mu(P)/2k$ and for all $p' \in P'$, $\gamma_L(p',C) \geq (1 + 2/k)\gamma_L(P,C)$,

(b) there is an i, $0 \leq i \leq \log^*(k)$ such that if $P'' = \{p \in P' : [1 + \log^{(i+1)}(k)/k]\gamma_L(P,C) \leq \gamma_L(p,C) < [1 + \log^{(i)}(k)/k]\gamma_L(P,C)\}$ then $\mu(P'') > \mu(P)/(2\log^{(i)}(k))^2$, where $\log^{(i)}$ is the i-fold iteration of log.

Proof Part (a): Because the contribution to $\varphi_L(P,C)$ is about the same for each path after Step 4a, the contribution to $\varphi_L(P,C)$ by any subset of P is approximately proportional to the size of the subset. A factor of $\lambda \geq 1 - cn^{-1/10}$, for c some constant, accounts for the the slight discrepancies in contribution to $\varphi_L(P,C)$. Let $\alpha = \frac{\mu(P')}{\mu(P)}$ be the fraction of paths retained in Step 4b. α is smallest when the paths retained have perfect quality and the paths eliminated in both Step 3 and Step 4b have quality just less than the thresholds for each of those steps. Also, we know the average quality of retained paths is at most $1 \leq 2\gamma_L(P,C)$, since $\gamma_L(P,C) > 1/2$. The derivation of a lower bound on α starts with this worst-case relationship.

$$\begin{aligned}
\gamma_L(P,C)\lambda \leq\ & \frac{1}{4}(1 - \frac{8}{k})\gamma_L(P,C) \\
& + (\frac{3}{4} - \alpha)(1 + \frac{2}{k})\gamma_L(P,C) \\
& + 2\alpha\gamma_L(P,C).
\end{aligned}$$

Collecting like terms and dividing by $\gamma_L(P,C)$, we have

$$\lambda \leq 1 - 2/k + \alpha(1 - 2/k)$$

and therefore $\alpha \geq 1/2k$. □

Note that at this point we are no longer concerned with decreases in γ; we have achieved a high minimum per-path quality. The remaining substeps of Step 4 make sure that all reduced input path-coloring pairs will contain exactly one bi-chromatic edge and will originate from an initial path-coloring pair that is valid and of high quality.

Step 4c anticipates condition 1 of Lemma 4 by eliminating those paths whose left subpath is not the left subpath of sufficiently many paths. Furthermore, the density of the stems is high following this step.

Lemma 9 Let $\mu(P) \geq 3n^{-1/100}$ be the path set density quality just prior to Step 4c of the lower bound strategy, and $|X^P|/|X|$ be the subsequent density of the stems. Then, $|X^P|/|X| \geq \mu(P)/2$.

Proof By definition, $|P_n^l| \geq (1 - l^2/n)|X||Y| \geq (1 - n^{-1/2})|X||Y|$, where the factor accounts for the possibility of common vertices in X and Y. Therefore, before Step 4c, $|P| \geq \mu(P)(1 - n^{-1/2})|X||Y|$. Since each stem not in $|X^P|$ contributes at most $n^{-1/100}|Y|$ paths to P, the total number of removed paths is at most $n^{-1/100}|X||Y|$. The total number of remaining paths is less than $|X^P||Y|$ so we have

$$|X^P||Y| + n^{-1/100}|X||Y| \geq \mu(P)(1 - n^{-1/2})|X||Y|.$$

Collecting terms and dividing by $|X||Y|$,

$$\frac{|X^P|}{|X|} \geq \mu(P)(1 - n^{-1/2}) - n^{-1/100} \geq \frac{\mu(P)}{2}$$

by the conditions on $\mu(P)$. □

Lemma 10 Let P and C be the path and coloring sets just after Step 4c of the lower bound strategy, where $\mu(C) \geq 2^{-(1/2)n^{1/100}}$ and $\gamma_L(P,C) > 1/2$. Furthermore, suppose $3 \leq l_L \leq (1/2)n^{1/100}$ is the left path length. Then, for all $p \in P$, $\gamma_L(p,C)\varphi_L(p,C) \geq |C_{n-l_R}|2^{-n^{1/100}}$.

Proof By definition for all paths $p \in P_n^l$, $\varphi_L(p,C_n) = l_L|C_n|/2^l = l_L|C_{n-l_R}|/2^{l_L}$. After Step 4a, every path $p \in P$ has

$$\begin{aligned}\varphi_L(p,C) &\geq \varphi_L(p,C_n)\mu(C)(1 - O(n^{-\frac{1}{10}})) \\ &= \frac{l_L\mu(C)|C_{n-l_R}|(1 - O(n^{-\frac{1}{10}}))}{2^{l_L}} \\ &> 2^{-n^{-1/100}+1}|C_{n-l_R}|\end{aligned}$$

by the bounds on $\mu(C)$ and l_L. Since $\gamma_L(p,C) \geq 1/2$, the desired result follows. □

The lower bound strategy is now prepared to reduce the input domain, path length, and the path and coloring sets so that all valid input pairs contain a bi-chromatic edge and are parts of valid, high quality original input pairs. The following lemma ensures that we can do this necessary clean-up without too much loss in input set densities or quality.

Lemma 11 Let P be the set of paths with distinct left stems X^P, $|X^P|/|X| \geq 20n^{-1/100}$, and C be the set of colorings prior to Step 4d of the lower bound strategy such that for all $p \in P$, $\gamma_L(p,C) \geq \Gamma$. Let $\mathcal{D}'$, P', C', $\gamma(P',C')$, and l' be the subsequent vertex domain, path and coloring sets, quality, and path length following Step 4d. Then there is a $\lambda \geq (1 - O(n^{-\frac{1}{10}}))$ such that, $n' = |\mathcal{D}'| = n - \sqrt{n}$, $l' = l_L$, $\mu'(P') = |P'|/|P_{n'}^{l'}| \geq |X^P|/(10|X|)$, $\mu'(C') = |C'|/|C_{n'}| \geq \mu(C)\lambda$, $\gamma(P',C') \geq \Gamma\lambda$.

Proof By Step 4a all paths $p \in P$ have $\varphi_L(p,C) \geq l_L|C_n|\mu(C)\lambda'/2^l$, where $\lambda' \geq 1 - cn^{-\frac{1}{10}}$, for some constant c. Because of the conditions guaranteed by Lemmas 9 and 10, we can apply Lemma 4 to get $\mathcal{D}'$, P'', C'', and $\lambda' = (1 - O(n^{-\frac{1}{10}}))$, such that $|P''|/|P_{n'}^{l'}| \geq |X^P|/(5|X|)$, $\mu(C)/\lambda \geq |C''|/|C_{n'}| \geq \mu(C)\lambda$, for all $p \in P''$

$$\begin{aligned}\gamma_L(p,C'')\varphi_L(p,C'') &\geq \Gamma\frac{l_L|C_n|\mu(C)\lambda'}{2^l}\frac{|C_{n'}|\lambda}{|C_{n-l_R}|} \\ &= \Gamma l_L|C_{n'}|\mu(C)\lambda\lambda'/2^{l_L},\end{aligned}$$

all colorings in C'' agree on $\mathcal{D} \setminus \mathcal{D}'$, and for each $x \in X^P$ there is a unique $y \in Y$ such that $(x;y) \in P''$, all vertices in this y are in $\mathcal{D} \setminus \mathcal{D}'$ and are colored the same as t.

Now we let P''' be the set of stems of paths in P'', and C' be the set of colorings of C'' restricted to

the set $\mathcal{D}'$. We clearly have $\mu'(P''') \geq |X^P|/(5|X|)$, $\mu(C)/\lambda \geq \mu'(C') \geq \mu(C)\lambda$ and for all $p \in P'''$,

$$\begin{aligned}\gamma(p,C')\varphi(p,C') &\geq \Gamma l'|C_{n'}|\mu(C)\lambda\lambda'/2^{l'} \\ &\geq \Gamma l'|C_{n'}|\mu(C')\lambda^2\lambda'/2^{l'}.\end{aligned}$$

We now let P' be those paths $p \in P'''$ such that $\varphi(p,C') = (1 \pm O(n^{-\frac{1}{10}}))l'\mu(C')|C_{n'}|/2^{l'}$, analogous to Steps 4a and 6. Then according to Lemma 7, $\mu(P') \geq \mu(P''')/2 \geq |X^P|/(10|X|)$, and we derive that for all paths $p \in P'$, $\gamma(p,C) \geq \Gamma\lambda''$, for some $\lambda'' = (1 - O(n^{-\frac{1}{10}}))$. □

We now summarize the effects of the lower bound strategy using only the (a) parts of the two-part lemmas. The next three lemmas progressively recapitulate the effects of quality-increasing loop, the round, and finally the entire lower bound strategy. The following lemma summarizes the effects of Step 4 on $\mu(P)$, $\mu(C)$, $\gamma(P,C)$, and l.

Lemma 12 Let $\mu(P) \geq 160kn^{-1/100}$, $\mu(C) \geq 2^{-(1/2)n^{1/100}}$, $\gamma(P,C) \geq 1/[2(1 - 2/k)]$, and $(1/2)n^{1/100} \geq l \geq 12k$ be the path and coloring set densities, protocol quality, and path length just prior to an iteration of Step 4 of the lower bound strategy. Let $\mu(P')$, $\mu(C')$, $\gamma(P',C')$, and l' be the corresponding values at the end of the iteration. Then, $\mu(P') \geq (1/80k)\mu(P)$, $\mu(C') \geq \mu(C)/2$, $\gamma(P',C') \geq (1 + 1/2k)\gamma(P,C)$, and $l' \geq l/4k$.

Proof By the previous lemmas, $\mu(P)$ declines by factors of at most 2 in each of Steps 4a and 4c, at most $2k$ in Step 4b, and at most 10 in Step 4d. $\mu(C)$ declines only by an $(1 - O(n^{-\frac{1}{10}}))$ amount and only in Step 4d. l declines by at most a $4k$ factor in Step 4d. Finally, $\gamma(P,C)$ increases by a $(1 + 2/k)$ factor in Step 4b and declines by a $(1-1/k)(1-O(n^{-\frac{1}{10}}))$ factor in Step 4d (since this is where we actually restrict to the left side quality) and only a $(1 - O(n^{-\frac{1}{10}}))$ factor in Step 4a. □

Once the strategy is out of the loop of Step 4, set density can be substantially increased by a path reduction to one side or the other in Steps 5 and 7, which includes an input clean-up identical to that of Lemma 11. We get a net improvement in path set density from $\mu(P)$ to $\Theta(\sqrt{\mu(P)})$. Thus we have:

Lemma 13 Let $\mu(P) \geq 2(80k)^k 2^{3\beta} n^{-1/100}$, $\mu(C) \geq 2^{3\beta+2k+1-(1/2)n^{1/100}}$, and $\gamma(P,C) \geq 1/[2(1 - 10/k)]$ be the path and coloring set densities and protocol quality just prior to a round of the lower bound strategy. Also, let $(1/2)n^{1/100} \geq l \geq 3(4k)^{2k+1}$ be the length of the paths at this point. If $\mu(P')$, $\mu(C')$, $\gamma(P',C')$, and l' are the corresponding values at the end of the round, assuming that x iterations through Step 4 were made, then $x \leq 2k$, $\mu(P') \geq (1/40)(1/80k)^{x/2}\sqrt{\mu(P)2^{-3\beta}}$, $\mu(C') \geq 2^{-3\beta-x-1}\mu(C)$, $\gamma(P',C') \geq (1 - 10/k)(1 + 1/(2k))^x\gamma(P,C)$, $l' \geq (4k)^{-x-1}l$.

Proof Since $\gamma(P,C)$ is greater than $1/2$ just prior to Step 4, increases by a factor of $(1 + 1/(2k))$ per iteration, and cannot exceed 1, we can use the fact that for all α, $0 \leq \alpha \leq 1$, $1 + \alpha \geq 2^\alpha$ to show that the maximum number, x, of iterations of Step 4 in a round is $2k$.

By the previous lemmas, $\mu(P)$ first declines by a factor of $2^{-3\beta}$ in Step 1, a factor of $1/80k$ for each iteration of Step 4, and a factor of 2 in the first part of Step 5, where $P \leftarrow P_L \bigcap P_R$. The density-increasing part of Step 5 then increases $\mu(P)$ to the square root of half this intermediate value. Steps 6 and 7 somewhat reduce this increased value of $\mu(P)$ by additional factors of 2 and 10.

In each round $\mu(C)$ declines by a factor of $2^{-3\beta}$ in Step 1 and by easily no more than a factor of 2 both for each iteration of Step 4 and for Step 7.

$\gamma(P,C)$ declines by a factor of $(1-1/k)$ in Step 2, a single factor of $(1 - 8/k)$ due to Steps 3 and 4b, and a factor of $(1 - O(n^{-\frac{1}{10}}))$ each for Steps 6 and 4d—for a total decline per round of at most a $(1 - 10/k)$ factor. From Lemma 12 we know $\gamma(P,C)$ increases by a factor of $(1 + 1/2k)$ per iteration of Step 4.

The path length l is shortened by a factor of $4k$ both for each iteration of Step 4 and for Step 7. □

We are now in a position to determine the number β of bits sent per round, constraints on the length l of the paths, and the maximum number t of rounds possible according to the lower bound strategy. In the following lemma we are able to obtain a bound on the total number of times the quality-increasing loop of Step 4 is executed, which allows us to bound the total decrease in path length in terms of t and k.

Lemma 14 Let $\beta \leq \frac{1}{300}\log n$, $n^{1/200} \leq l_0 < (1/2)n^{1/100}$, where l_0 is the initial path length, $40 \leq k \leq \frac{1}{4000}\log n/\log\log n$, and the initial protocol quality be 3/4. Then, for $t \leq k/35$, after t rounds of the strategy, if Step 4 is iterated a total of y times, the following hold: $\mu(P) \geq n^{-1/200}$, $\mu(C) \geq 2^{-3\beta t-y-t}$, $\gamma(P,C) \geq (1-10/k)^t(1+1/2k)^y(3/4)$, $l \geq (4k)^{-t-y}l_0$, and $y \leq 40t+2k$.

Proof By induction on t. First note that $\gamma(P,C) \geq (1-10/k)^t(3/4) \geq 2^{-\frac{5}{3}(10t/k)}(3/4) \geq .51$ since $1-\alpha \geq 2^{-\frac{5}{3}\alpha}$ for $0 \leq \alpha \leq 1/4$ and since t is at most $k/35$. It is easy to verify that all the conditions of Lemma 13 are satisfied. Suppose that Step 4 is iterated $y'-y$ times in the round. Then $y'-y \leq 2k$, and at the end of the round it is easy to see that $\mu(P) \geq n^{-1/200}$, $\mu(C) \geq 2^{-(3\beta+1)(t+1)-y'}$, $\gamma(P,C) \geq (1-10/k)^{t+1}(1+1/(2k))^{y'}(3/4)$, and $l \geq (2k)^{-2(t+1)-2y'}l_0$. Since $\gamma(P,C) \leq 1$, we can derive $y' \leq 40(t+1)+2k$. □

Let k and β be chosen as large as possible in Lemma 14 and consider any $t \leq k/35$. After t rounds, the strategy is correct on at least a .51 fraction of the inputs remaining but, for n sufficiently large, l is sufficiently large, and by Corollary 2 the function takes on any fixed value on less than this fraction of the inputs so the protocol must continue. Thus the total number of bits communicated must be more than $t\beta = \Omega(\log^2 n/\log\log n)$.

In the calculations of Lemmas 12, 13, and 14, the most important factors influencing the value of k are the decrease in path length by a factor of $4k$ in Steps 2 and 4e, and the decrease in path set density by a factor of $2k$ in Step 4b. By applying part (b) of Lemma 6 we can show that the total decrease in path length over the course of the strategy is at most $2^{O(k)}$ rather than $k^{O(k)}$. By using the threshold from part (b) of Lemma 8 in Step 4b we can show that the total decrease in path set density through the iterations of Step 4 in a single round is also $2^{O(k)}$ rather than $k^{O(k)}$. Thus we can show the following:

Theorem 15 For $0 \leq \epsilon < 1/2$, the ϵ-error randomized communication complexity of the path-coloring decision problem on a graph with n nodes is $\Omega(\log^2 n)$.

Acknowledgments We wish to thank Avi Wigderson for suggesting the path-coloring problem as a function that might separate randomized from nondeterministic communication complexity, as well as Johan Håstad for his helpful comments.

References

[AUY83] A. V. Aho, J. D. Ullman, and M. Yannakakis. On notions of information transfer in VLSI circuits. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, pages 133–139, Boston, MA, April 1983.

[BI87] M. Blum and R. Impagliazzo. Generic oracles and oracle classes. In *28th Annual Symposium on Foundations of Computer Science*, pages 118–126, Los Angeles, CA, October 1987. IEEE.

[DF89] D. Dolev and T. Feder. Multiparty communication complexity. In *30th Annual Symposium on Foundations of Computer Science*, pages 428–433, Research Triangle Park, NC, October 1989. IEEE. Corrected version.

[Für87] M. Fürer. The power of randomness for communication complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, pages 178–181, New York, NY, May 1987.

[GH89] M. Goldmann and J. Håstad. A lower bound for monotone clique using a communication game. Manuscript, 1989.

[HH87] J. Hartmanis and L. Hemachandra. One-way functions, robustness, and non-isomorphism of NP-complete sets. Technical Report DCS TR86-796, Cornell University, 1987.

[HR88] Bernd Halstenberg and Rüdiger Reischuk. On different modes of communication. In *Proceedings of the Twentieth*

Annual ACM Symposium on Theory of Computing, pages 162–172, Chicago, IL, May 1988.

[KW88] M. Karchmer and A. Wigderson. Monotone circuits for connectivity require super-logarithmic depth. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, pages 539–550, Chicago, IL, May 1988.

[Law92] Joan Lawry. PhD thesis, University of Washington, 1992.

[LNNW91] L. Lovász, M. Naor, I. Newman, and A. Wigderson. Search problems in the decision tree model. In *32nd Annual Symposium on Foundations of Computer Science*, pages 576–585, San Juan, Puerto Rico, October 1991. IEEE.

[MS82] K. Mehlhorn and E. M. Schmidt. Las Vegas is better than determinism in VLSI and distributed computing. In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*, pages 330–337, San Francisco, CA, May 1982.

[Raz88] A. A. Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. Manuscript, 1988.

[RW89] Ran Raz and Avi Wigderson. Probabilistic communication complexity of Boolean relations. In *30th Annual Symposium on Foundations of Computer Science*, pages 562–567, Research Triangle Park, NC, October 1989. IEEE. Full version.

[RW90] Ran Raz and Avi Wigderson. Monotone circuits for matching require linear depth. In *Proceedings of the Twenty Second Annual ACM Symposium on Theory of Computing*, pages 287–292, Baltimore, MD, May 1990.

[Sni85] M. Snir. Lower bounds on probabilistic linear decision trees. *Theoretical Computer Science*, 38:69–82, 1985.

[SW86] M. Saks and A. Wigderson. Probabilistic Boolean decision trees and the complexity of evaluating games trees. In *27th Annual Symposium on Foundations of Computer Science*, pages 29–38, Toronto, Ontario, October 1986. IEEE.

[Tar88] G. Tardos. Query complexity, or why is it difficult to separate $NP^A \bigcap coNP^A$ by a random oracle A? Manuscript, 1988.

[Yao79] A. C. Yao. Some complexity questions related to distributive computing. In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, pages 209–213, Atlanta, GA, April-May 1979.

[Yao81] A. C. Yao. The entropic limitations of VLSI computations. In *Proceedings of the Thirteenth Annual ACM Symposium on Theory of Computing*, pages 308–311, Milwaukee, WI, May 1981.

[Yao83] A. C. Yao. Lower bounds by probabilistic arguments. In *24th Annual Symposium on Foundations of Computer Science*, pages 420–428, Tucson, AZ, November 1983. IEEE.