Brief Announcement: An Exponential Separation Between Randomized and Deterministic Complexity in the LOCAL Model*

Yi-Jun Chang
University of Michigan
Ann Arbor, MI
cyjjun@umich.edu

Tsvi Kopelowitz
University of Michigan
Ann Arbor, MI
tsvi@umich.edu

Seth Pettie
University of Michigan
Ann Arbor, MI
pettie@umich.edu

ABSTRACT

Over the past 30 years numerous algorithms have been designed for symmetry breaking problems in the LOCAL model, such as maximal matching, MIS, vertex coloring, and edge-coloring. For most problems the best randomized algorithm is at least exponentially faster than the best deterministic algorithm. In this paper we prove that these exponential gaps are *necessary* and establish numerous connections between the deterministic and randomized complexities in the LOCAL model. Each of our results has a very compelling take-away message:

1. **Fast Δ -coloring of trees requires random bits.** Building on the recent randomized lower bounds of Brandt et al. [6], we prove that the randomized complexity of Δ -coloring a tree with maximum degree Δ is $\Theta(\log_{\Delta} \log n)$, for any $\Delta \geq 55$, whereas its deterministic complexity is $\Theta(\log_{\Delta} n)$ for any $\Delta \geq 3$.¹ This also establishes a large separation between the deterministic complexity of Δ -coloring and $(\Delta + 1)$ -coloring trees.
2. **Randomized lower bounds imply deterministic lower bounds.** We prove that any deterministic algorithm for a natural class of problems that runs in $O(1) + o(\log_{\Delta} n)$ rounds can be transformed to run in $O(\log^* n - \log^* \Delta + 1)$ rounds. If the transformed algorithm violates a lower bound (even allowing randomization), then one can conclude that the problem requires $\Omega(\log_{\Delta} n)$ time deterministically. (This gives an alternate proof that deterministically Δ -coloring a tree with small Δ takes $\Omega(\log_{\Delta} n)$ rounds.)

*A more detailed version of this paper appears in [7]. This work is supported by NSF grants CCF-1217338, CNS-1318294, and CCF-1514383.

¹For simplicity, here we suppress any $\log^* n$ additive term in $\Theta(\cdot)$.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

PODC'16 July 25-28, 2016, Chicago, IL, USA

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-3964-3/16/07.

DOI: <http://dx.doi.org/10.1145/2933057.2933079>

3. **Deterministic lower bounds imply randomized lower bounds.** We prove that the randomized complexity of any natural problem on instances of size n is at least its deterministic complexity on instances of size $\sqrt{\log n}$. This shows that a *deterministic* $\Omega(\log_{\Delta} n)$ lower bound for any problem (Δ -coloring a tree, for example) implies a *randomized* $\Omega(\log_{\Delta} \log n)$ lower bound. It also illustrates that the *graph shattering* technique employed in recent randomized symmetry breaking algorithms is *absolutely essential* to the LOCAL model. For example, it is provably impossible to improve the $2^{O(\sqrt{\log \log n})}$ terms in the complexities of the best MIS and $(\Delta + 1)$ -coloring algorithms without *also* improving the $2^{O(\sqrt{\log n})}$ -round Panconesi-Srinivasan algorithms.

Keywords

Coloring; distributed algorithm; symmetry breaking

1. INTRODUCTION

One of the central problems of theoretical computer science is to determine the value of *random bits*. If the distinction is between computable vs. uncomputable functions, random bits are provably useless in centralized models (Turing machines). However, this is not true in the distributed world! The celebrated Fischer-Lynch-Patterson theorem [10] states that asynchronous deterministic agreement is impossible with one unannounced failure, yet it is possible to accomplish with probability 1 using randomization [4].

In this paper we examine the value of random bits in Linial's [16] LOCAL model, which, for the sake of clarity, we bifurcate into two models RandLOCAL and DetLOCAL. In both models the input graph $G = (V, E)$ and communications network are identical. Each vertex hosts a processor and all vertices run the same algorithm. Each edge supports communication in both directions. The computation proceeds in synchronized *rounds*. In a round, each processor performs some computation and sends a message along each incident edge, which is delivered before the beginning of the next round. Each vertex v is initially aware of its degree $\deg(v)$ and certain global parameters such as $n \stackrel{\text{def}}{=} |V|$, $\Delta = \Delta(G) \stackrel{\text{def}}{=} \max_{v \in V} \deg(v)$, and possibly others.² In the LOCAL model the only measure of efficiency is the number

²The assumption that global parameters are common knowledge can sometimes be removed; see Korman, Sereni, and Viennot [14].

of rounds. *All local computation is free and the size of messages is unbounded.* Henceforth “time” refers to the number of rounds.

DetLOCAL: In order to avoid trivial impossibilities, all vertices are assumed to hold unique $\Theta(\log n)$ -bit IDs. Except for the registers holding $\deg(v)$ and $\text{ID}(v)$, the initial state of v is identical to every other vertex. The algorithm executed at each vertex is deterministic.

RandLOCAL: In this model each vertex may locally generate an unbounded number of independent truly random bits. (There are no globally shared random bits.) Except for the register holding $\deg(v)$, the initial state of v is identical to every other vertex. Algorithms in this model operate for a specified number of rounds and have some probability of *failure*, the definition of which is problem specific. We usually only consider algorithms whose global probability of failure is at most $1/\text{poly}(n)$.

Observe that the lack of IDs in RandLOCAL is not a practical limitation. Before the first round each node can locally generate a random $O(\log n)$ -bit ID, which is unique with probability $1/\text{poly}(n)$.

Early work in the LOCAL models suggested that randomness is of limited help. Naor [17] showed that Linial’s $\Omega(\log^* n)$ lower bound [16] for 3-coloring the ring holds even in RandLOCAL, and Naor and Stockmeyer [18] proved that the class of problems solvable by $O(1)$ -round algorithms is the same in RandLOCAL and DetLOCAL. However, in the intervening decades we have seen dozens of examples of symmetry breaking algorithms for RandLOCAL that are substantially faster than their counterparts in DetLOCAL; see [3] for a survey of three archetypal problems: maximal independent set (MIS), maximal matching, and $(\Delta + 1)$ -coloring.

Graph Shattering.

Randomized symmetry breaking algorithms are typically exponentially faster *in two ways*. Their dependence on Δ is exponentially faster and their dependence on n is usually identical to the best deterministic complexity, but for $\text{poly}(\log n)$ -size instances, for example, $2^{O(\sqrt{\log n})}$ becomes $2^{O(\sqrt{\log \log n})}$. This second phenomenon is no coincidence! It is a direct result of the *graph shattering* approach to symmetry breaking used in [3] and further in [11, 13]. The idea is to apply some randomized procedure that fixes some fragment of the output (e.g., part of the MIS is fixed, part of the coloring is fixed, etc.), thereby effectively removing a large fraction of the vertices from further consideration. If it can be shown that the connected components in the subgraph still under consideration have size $\text{poly}(\log n)$, one can revert to the best available *deterministic* algorithm and solve the problem on each component of the “shattered” graph in parallel.

Lower Bounds in the LOCAL Model.

Until recently, the main principle used to prove lower bounds in the LOCAL model was *indistinguishability*. The first application of this principle was by Linial [16] himself, who argued that any algorithm for coloring degree- Δ trees either uses $\Omega(\Delta/\log \Delta)$ colors or takes $\Omega(\log \Delta n)$ time. The proof is as follows (i) in $o(\log \Delta n)$ time, a vertex *cannot always distinguish* whether the input graph G is a tree or a

graph with girth $\Omega(\log \Delta n)$, (ii) for all Δ and all n , there exists a degree- Δ graph with girth $\Omega(\log \Delta n)$ and chromatic number $\chi = \Omega(\Delta/\log \Delta)$, hence³ (iii) any $o(\log \Delta n)$ -time algorithm for coloring trees could also color such a graph, and therefore must use at least χ colors. A more subtle indistinguishability argument was used by Kuhn, Moscibroda, and Wattenhofer [15], who showed that $O(1)$ -approximate vertex cover, maximal matching, MIS, and several other problems have $\Omega(\min\{\log \Delta/\log \log \Delta, \sqrt{\log n/\log \log n}\})$ lower bounds. Recently, Bar-Yehuda, Censor-Hillel, and Schwartzman [1] showed that a $(2 + \epsilon)$ -approximate vertex cover can be found in $O(\log \Delta/\epsilon \log \log \Delta)$ time, matching the above lower bound. By its nature, indistinguishability is not very good at separating randomized and deterministic complexities. Recently, Brandt et al. [6] developed a lower bound technique that explicitly incorporates error probabilities and proved that several problems on graphs with constant Δ take $\Omega(\log \log n)$ time in RandLOCAL (with error probability $1/\text{poly}(n)$) such as *sinkless orientation*, *sinkless coloring*, and Δ -coloring. Since the existence of a sinkless orientation can be proved with the Lovász local lemma (LLL), this gave $\Omega(\log \log n)$ lower bounds on distributed algorithms for the constructive LLL. See [8, 11] for upper bounds on the distributed LLL.

2. NEW RESULTS

We exhibit an exponential separation between RandLOCAL and DetLOCAL for several *specific* symmetry breaking problems. More generally, we give new connections between the randomized and deterministic complexities of all *locally checkable labeling* (LCL) problems, a class that includes essentially any natural symmetry breaking problem.

Separation of RandLOCAL and DetLOCAL.

We extend Brandt et al.’s [6] randomized lower bound as follows: on degree- Δ graphs, Δ -coloring takes $\Omega(\log \Delta \log n)$ time in RandLOCAL and $\Omega(\log \Delta n)$ time in DetLOCAL. The hard graphs in this lower bound have girth $\Omega(\log \Delta n)$, so by the indistinguishability principle, these lower bounds also apply to Δ -coloring trees. On the upper bound side, Barenboim and Elkin [2] showed that for $\Delta \geq 3$, Δ -coloring trees takes $O(\log \Delta n + \log^* n)$ time in DetLOCAL. We give an elementary proof that for $\Delta \geq 55$, Δ -coloring trees can be done in $O(\log \Delta \log n + \log^* n)$ time in RandLOCAL, matching Brandt et al.’s [6] lower bound up to the $\log^* n$. A more complicated algorithm for Δ -coloring trees could be derived from [20], for $\Delta > \Delta_0$ and some very large constant Δ_0 .⁴

Randomized lower bounds imply deterministic lower bounds.

We give a second, more generic proof that Δ -coloring trees

³Linial [16] actually only used the existence of Δ -regular graphs with high girth and chromatic number $\Omega(\sqrt{\Delta})$. See [5] for constructions with chromatic number $\Omega(\Delta/\log \Delta)$.

⁴The reason we are interested in minimizing the $\Delta_0 \leq \Delta$ for which the algorithm works is somewhat technical. It seems as if Δ -coloring trees is a problem whose character makes a *qualitative* transition when Δ is a small enough constant. Using our technique (graph shattering) we may be able to replace 55 with a smaller constant, *but not too small*. Any algorithm that 3-colors 3-regular trees, for example, will need to be qualitatively very different in its design.

takes $\Omega(\log_{\Delta} n)$ time. The proof shows that any $f(\Delta) + o(\log_{\Delta} n)$ time algorithm for an LCL problem can be transformed in a black box way to run in $O((1 + f(\Delta))(\log^* n - \log^* \Delta + 1))$ time. Thus, on bounded-degree graphs, there are no “natural” deterministic time bounds between $\omega(\log^* n)$ and $o(\log n)$. Any $\omega(\log^* n)$ lower bound for bounded degree graphs (in either RandLOCAL or DetLOCAL) immediately implies an $\Omega(\log n)$ lower bound in DetLOCAL. This reduction can be parameterized in many different ways. Under a different parametrization it shows that any $O(\log^{1-\frac{1}{k+1}} n)$ -time DetLOCAL algorithm for an LCL problem can be transformed to run in $O(\log^k \Delta (\log^* n - \log^* \Delta + 1))$ -time. For example, if one were to develop a deterministic $O(\sqrt{\log n})$ -time MIS or maximal matching algorithm—almost matching one of the KMW [15] lower bounds—it would immediately imply an $O(\log \Delta (\log^* n - \log^* \Delta + 1))$ -time MIS/maximal matching algorithm, which almost matches the other KMW lower bound, for $\Delta > \log^{O(1)} n$. By some strange coincidence, [3] gave an analogous reduction for MIS/maximal matching in bounded arboricity graphs, but for RandLOCAL and in the reverse direction. Specifically, any $O(\log^k \Delta + f(n))$ -time RandLOCAL MIS/maximal matching algorithm can be transformed into an $O(\log^{1-\frac{1}{k+1}} n + f(n))$ -time RandLOCAL algorithm for bounded arboricity graphs.

Deterministic lower bounds imply randomized lower bounds.

We prove that for any LCL problem, its RandLOCAL complexity on instances of size n is at least its DetLOCAL complexity on instances of size $\sqrt{\log n}$. This reverses the implication proved above. For example, if we begin with a proof that Δ -coloring takes $\Omega(\log_{\Delta} n)$ time in DetLOCAL, then we conclude that it must take $\Omega(\log_{\Delta} \log n)$ time in RandLOCAL. This result has a very clear take-away message: the *graph shattering* technique applied by recent randomized symmetry breaking algorithms [3, 11] is *inherent* to the RandLOCAL model and every optimal RandLOCAL algorithm for instances of size n must, in some way, encode an optimal DetLOCAL algorithm on $\text{poly}(\log n)$ -size instances. It is therefore impossible to improve the $2^{O(\sqrt{\log \log n})}$ terms in the RandLOCAL MIS and coloring algorithms of [3, 11, 13, 9] without also improving the $2^{O(\sqrt{\log n})}$ -time DetLOCAL algorithms of Panconesi and Srinivasan [19], and it is impossible to improve the $O(\log^4 \log n)$ term in the RandLOCAL maximal matching algorithm of [3] without also improving the $O(\log^4 n)$ DetLOCAL maximal matching algorithm of [12].

3. REFERENCES

- [1] R. Bar-Yehuda, K. Censor-Hillel, and G. Schwartzman. A distributed $(2 + \epsilon)$ -approximation for vertex cover in $O(\log \Delta / \epsilon \log \log \Delta)$ rounds. In *Proceedings 35th Annual ACM Symposium on Principles of Distributed Computing (PODC)*, 2016.
- [2] L. Barenboim and M. Elkin. Sublogarithmic distributed MIS algorithm for sparse graphs using Nash-Williams decomposition. *Distributed Computing*, 22(5-6):363–379, 2010.
- [3] L. Barenboim, M. Elkin, S. Pettie, and J. Schneider. The locality of distributed symmetry breaking. *J. ACM*, 2016. to appear.
- [4] M. Ben-Or. Another advantage of free choice: Completely asynchronous agreement protocols. In *Proceedings of the 2nd ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC)*, pages 27–30, 1983.
- [5] B. Bollobás. Chromatic number, girth and maximal degree. *Discrete Mathematics*, 24(3):311–314, 1978.
- [6] S. Brandt, O. Fischer, J. Hirvonen, B. Keller, T. Lempinen, J. Rybicki, J. Suomela, and J. Uitto. A lower bound for the distributed Lovász local lemma. In *Proceedings 48th ACM Symposium on the Theory of Computing (STOC)*, 2016.
- [7] Y.-J. Chang, T. Kopelowitz, and S. Pettie. An exponential separation between randomized and deterministic complexity in the local model. *CoRR*, abs/1602.08166, 2016.
- [8] K.-M. Chung, S. Pettie, and H.-H. Su. Distributed algorithms for the Lovász local lemma and graph coloring. In *Proceedings 33rd ACM Symposium on Principles of Distributed Computing (PODC)*, pages 134–143, 2014.
- [9] M. Elkin, S. Pettie, and H. H. Su. $(2\Delta - 1)$ -edge coloring is much easier than maximal matching in the distributed setting. In *Proceedings 26th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 355–370, 2015.
- [10] M. J. Fischer, N. A. Lynch, and M. Paterson. Impossibility of distributed consensus with one faulty process. *J. ACM*, 32(2):374–382, 1985.
- [11] M. Ghaffari. An improved distributed algorithm for maximal independent set. In *Proceedings 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 270–277, 2016.
- [12] M. Hańćkowiak, M. Karoński, and A. Panconesi. On the distributed complexity of computing maximal matchings. *SIAM J. Discrete Mathematics*, 15(1):41–57 (electronic), 2001.
- [13] D. Harris, J. Schneider, and H.-H. Su. Distributed $(\Delta + 1)$ -coloring in sublogarithmic rounds. In *Proceedings 48th ACM Symposium on Theory of Computing (STOC)*, 2016.
- [14] A. Korman, J.-S. Sereni, and L. Viennot. Toward more localized local algorithms: removing assumptions concerning global knowledge. *Distributed Computing*, 26(5-6):289–308, 2013.
- [15] F. Kuhn, T. Moscibroda, and R. Wattenhofer. Local computation: Lower and upper bounds. *J. ACM*, 63(2):17:1–17:44, 2016.
- [16] N. Linial. Locality in distributed graph algorithms. *SIAM J. Comput.*, 21(1):193–201, 1992.
- [17] M. Naor. A lower bound on probabilistic algorithms for distributive ring coloring. *SIAM J. Discrete Mathematics*, 4(3):409–412, 1991.
- [18] M. Naor and L. J. Stockmeyer. What can be computed locally? *SIAM J. Comput.*, 24(6):1259–1277, 1995.
- [19] A. Panconesi and A. Srinivasan. On the complexity of distributed network decomposition. *J. Algor.*, 20(2):356–374, 1996.
- [20] S. Pettie and H.-H. Su. Distributed algorithms for coloring triangle-free graphs. *Information and Computation*, 243:263–280, 2015.