



HAL
open science

Online Locality Meets Distributed Quantum Computing

Amirreza Akbari, Xavier Coiteux-Roy, Francesco d'Amore, François Le Gall,
Henrik Lievonen, Darya Melnyk, Augusto Modanese, Shreyas Pai,
Marc-Olivier Renou, Václav Rozhoň, et al.

► **To cite this version:**

Amirreza Akbari, Xavier Coiteux-Roy, Francesco d'Amore, François Le Gall, Henrik Lievonen, et al..
Online Locality Meets Distributed Quantum Computing. 2024. hal-04490541v2

HAL Id: hal-04490541

<https://hal.science/hal-04490541v2>

Preprint submitted on 9 Apr 2024 (v2), last revised 5 Nov 2024 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Online Locality Meets Distributed Quantum Computing

Amirreza Akbari · Aalto University, Finland

Xavier Coiteux-Roy · School of Computation, Information and Technology, Technical University of Munich, Germany · Munich Center for Quantum Science and Technology, Germany

Francesco d'Amore · Aalto University, Finland · Bocconi University, Italy · BIDS, Italy

François Le Gall · Nagoya University, Japan

Henrik Lievonen · Aalto University, Finland

Darya Melnyk · Technische Universität Berlin, Germany

Augusto Modanese · Aalto University, Finland

Shreyas Pai · Aalto University, Finland

Marc-Olivier Renou · Inria, Université Paris-Saclay, Palaiseau, France · CPHT, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France

Václav Rozhoň · ETH Zurich, Switzerland

Jukka Suomela · Aalto University, Finland

Abstract. We extend the theory of locally checkable labeling problems (LCLs) from the classical LOCAL model to a number of other models that have been studied recently, including the quantum-LOCAL model, finitely-dependent processes, non-signaling model, dynamic-LOCAL model, and online-LOCAL model [e.g. STOC 2024, ICALP 2023].

First, we demonstrate the advantage that finitely-dependent processes have over the classical LOCAL model. We show that all LCL problems solvable with locality $O(\log^* n)$ in the LOCAL model admit a finitely-dependent distribution (with constant locality). In particular, this gives a finitely-dependent coloring for regular trees, answering an open question by Holroyd [2023]. This also introduces a new formal barrier for understanding the distributed quantum advantage: it is not possible to exclude quantum advantage for any LCL in the $\Theta(\log^* n)$ complexity class by using non-signaling arguments.

Second, we put limits on the capabilities of all of these models. To this end, we introduce a model called randomized online-LOCAL, which is strong enough to simulate e.g. SLOCAL and dynamic-LOCAL, and we show that it is also strong enough to simulate any non-signaling distribution and hence any quantum-LOCAL algorithm. We prove the following result for rooted trees: if we can solve an LCL problem with locality $o(\log \log n)$ in the randomized online-LOCAL model, we can solve it with locality $O(\log^* n)$ in the classical deterministic LOCAL model.

Put together, these results show that in rooted trees the set of LCLs that can be solved with locality $O(\log^* n)$ is the same across all these models: classical deterministic and randomized LOCAL, quantum-LOCAL, non-signaling model, dynamic-LOCAL, and deterministic and randomized online-LOCAL.

1 Introduction

In this work, we connect three distinct lines of research that have recently explored extensions of the classical LOCAL model of distributed computing:

1. Distributed quantum computing and non-signaling distributions [3, 26, 33].
2. Finitely-dependent processes [40, 41, 43].
3. Locality in online graph algorithms and dynamic graph algorithms [2, 21].

We prove new results on the capabilities and limitations of all of these models of computing, for *locally checkable labeling problems* (LCLs). Our work implies limitations on the *quantum advantage* in the distributed setting, and we also exhibit a new *barrier* for proving tighter bounds.

1.1 Classical models

Let us first recall the definitions of the classical models of distributed graph algorithms [50, 56] that form the foundation for our work; we will keep it brief here and postpone the formal definitions to Section 3:

- **Deterministic LOCAL:** Our input graph $G = (V, E)$ represents a computer network; each node $v \in V$ is a computer and each edge $\{u, v\} \in E$ is a communication link between two computers. Each node is labeled with a unique identifier from $\{1, 2, \dots, \text{poly}(|V|)\}$. All nodes follow the same distributed algorithm. Initially a node is only aware of its own identifier and its own degree. Computation proceeds in synchronous rounds, and in each round a node can send and receive a message to and from each neighbor and update its state. Eventually each node must stop and announce its local output (its part of the solution, e.g. in graph coloring its own color). The *running time*, *round complexity*, or *locality* of the algorithm is the (worst-case) number of rounds $T(n)$ until the algorithm stops in any n -node graph.
- **Randomized LOCAL:** As above, but each node has also access to a private source of random bits.

We also define the following variants (see e.g. [46] for discussion on the knowledge of global information):

- **Deterministic LOCAL (shared):** Deterministic LOCAL with shared global information. The set of nodes and their unique identifiers is globally known. In particular, we know the value of $n = |V|$.
- **Randomized LOCAL (shared):** Randomized LOCAL with shared global information and shared randomness. The set of nodes and their unique identifiers is globally known, and in addition to the private sources of random bits, there is also a shared source of random bits that all nodes can access.

We can interpret the shared versions of the models so that we get to *see* the set of nodes V and their unique identifiers in advance, and we can also *initialize* the nodes as we want based on this information (and hence in the randomized model, we can also initialize all nodes with the same shared random string), but the set of edges E is only revealed later. This interpretation will be useful especially with the quantum models.

1.2 Landscape of LCL problems

There has been more than three decades of work on understanding the capabilities and limitations of the classical deterministic and randomized LOCAL models, but for our purposes most interesting is the recent line of work that has studied distributed algorithms for *locally checkable labeling problems*, or LCLs. This is a family of problems first introduced by Naor and Stockmeyer [52]. LCL problems are graph problems that can be defined by specifying a *finite set of valid neighborhoods*—for example, the task of coloring graphs of maximum degree 5 with 7 colors is an example of an LCL problem.

Since 2016, we have seen a large body of work dedicated to understanding the computational complexity of LCL problems in the deterministic and randomized LOCAL models [4–6, 6–9, 11, 15, 16, 18, 19, 32,

34, 36, 37, 57], and nowadays there are even algorithms and computer tools available for exploring such questions [9, 22, 54]. As a result of this large international research effort, a landscape of the localities of LCL problems emerges [59]. We can classify LCL problems in discrete classes based on their locality, and we also understand how much randomness helps in comparison with deterministic algorithms.

Our main goal in this work is to extend this understanding of LCL problems far beyond the classical models, and especially explore what can be computed very fast in models that are much stronger than deterministic or randomized LOCAL.

1.3 Quantum-LOCAL and finitely-dependent processes

Let us start our exploration of stronger models with distributed quantum computation. The key question is understanding the *distributed quantum advantage*: what can we solve faster if our nodes are quantum computers and our edges are quantum communication channels? There is a long line of prior work that has explored variants of this theme in different models of distributed computing [3, 17, 26, 31, 33, 44, 45, 47, 48, 51, 61–63], but for our purposes these are the models of interest:

- **Quantum-LOCAL:** Our model of computing is similar to the deterministic LOCAL model above, but now with quantum computers and quantum communication links. More precisely, the quantum computers manipulate local states consisting of an unbounded number of qubits with arbitrary unitary transformations, the communication links are quantum communication channels (adjacent nodes can exchange any number of qubits), and the local output can be the result of any quantum measurement.
- **Quantum-LOCAL (shared):** Quantum-LOCAL with shared global information and a shared quantum state. As above, but now the algorithm may inspect and manipulate the set of nodes (before any edges are revealed). In particular, it may initialize the nodes with a globally shared entangled state.

As quantum theory intrinsically involves randomness, quantum-LOCAL is at least as strong as randomized LOCAL. There are some (artificial) problems that are known to be solvable much faster in quantum-LOCAL than deterministic or randomized LOCAL [48]; however, whether any LCL admits such a quantum advantage is a major open question in the field, and also one of the main motivations of our work.

1.3.1 Finitely-dependent processes and non-signaling model

Directly analyzing quantum-LOCAL is beyond the scope of current techniques. In essence, the only known technique for proving limitations of quantum-LOCAL is *sandwiching* it between the classical randomized-LOCAL model and more powerful models than quantum-LOCAL that do not explicitly refer to quantum information. These more powerful models are based on the *physical causality principle* (a.k.a. *non-signaling principle*). The idea is perhaps easiest to understand with the help of the following thought experiment:

Example 1.1. Fix a distributed algorithm \mathcal{A} in the quantum-LOCAL model (with shared global information and quantum state) that runs in T rounds on graphs with n nodes. Let $G = (V, E)$ be some n -node input graph. Apply \mathcal{A} repeatedly to G to obtain some probability distribution $Y(G)$ of outputs. Now fix some subset of nodes $U \subseteq V$, and consider the restriction of $Y(G)$ to U , in notation $Y(G) \upharpoonright_U$. Let $G[U, T]$ be the radius- T neighborhood of set U in G . Now modify G outside $G[U, T]$ to obtain a different n -node graph G' with $G[U, T] = G'[U, T]$. Apply \mathcal{A} to G' repeatedly, and we obtain another probability distribution $Y(G')$ of outputs. If $Y(G) \upharpoonright_U \neq Y(G') \upharpoonright_U$, it would be possible to use \mathcal{A} to transmit information in T time steps between two parties, Alice and Bob, that are within distance $T + 1$ from each other: Bob holds all nodes of U , and he can, therefore, observe $Y(G) \upharpoonright_U$, while Alice controls the graph outside $G[U, T] = G'[U, T]$, and she can, therefore, instantiate either G or G' . This would enable Alice to send a signal to Bob even if no physical communication occurred from Alice to Bob (as they are at distance $T + 1$ from each other and only T communication steps occurred), and thus violate the non-signaling principle.

This thought experiment suggests the following definition, also known as the φ -LOCAL model and the causal model [3, 33]:

- **Non-signaling model:** We can produce an arbitrary output distribution as long as it does not violate the non-signaling principle: if we fix any set of nodes U , modifying the structure of the input graph at more than a distance $T(n)$ from U does not affect the output distribution of U .

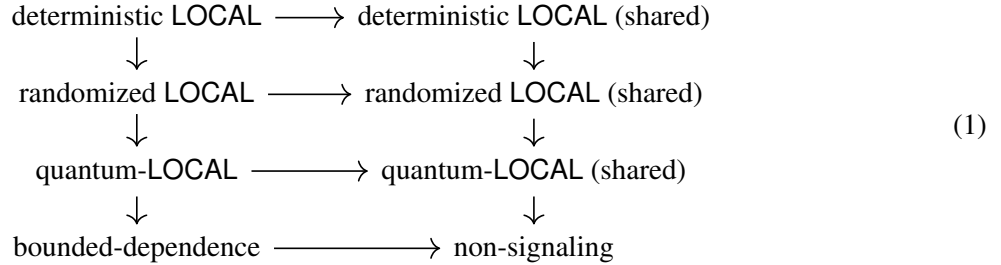
We also need to introduce the following definition to better connect our work with the study of finitely-dependent processes and in particular finitely-dependent colorings [40, 41, 43]:

- **Bounded-dependence model:** We can produce an arbitrary output distribution as long as it does not violate the non-signaling principle, and, furthermore, distant parts are independent: if we fix any sets of nodes U_1 and U_2 such that their radius- $T(n)$ neighborhoods are disjoint, then the output labels of U_1 are independent of the output labels of U_2 .

Now if we set $T(n) = O(1)$, algorithms in the bounded-dependence model are in essence what is usually called *finitely-dependent processes*.

1.3.2 Model hierarchy

Now we can connect all the above models with each others as follows, sandwiching the two versions of quantum-LOCAL between other models (see Appendix A for details; the connection with the non-signaling model is known [3, 33] but the connection with the bounded-dependence model is to our knowledge new):



Here an arrow $M_1 \rightarrow M_2$ indicates that an algorithm with locality (or round complexity) $T(n)$ in model M_1 implies an algorithm with the same locality in M_2 . For some problems it is possible to prove near-tight bounds for quantum-LOCAL by using diagram (1). For example, a very recent work [26] used these connections to prove limits for the distributed quantum advantage in approximate graph coloring: they prove an upper bound for the deterministic LOCAL model and a near-matching lower bound for the non-signaling model.

1.4 Contribution 1: symmetry breaking with finitely-dependent processes

Now we are ready to state our first contribution. Recall the following gap result by Chang et al. [20]: all LCL problems that can be solved with locality $o(\log n)$ in deterministic LOCAL or with locality $o(\log \log n)$ in randomized LOCAL can also be solved with locality $O(\log^* n)$ in deterministic LOCAL. The class of problems with locality $\Theta(\log^* n)$ contains in essence all *symmetry-breaking problems*: these are problems that could be solved with constant locality if only we had some means of breaking symmetry (e.g. distance- k coloring for some constant k would suffice). In Section 5 we show the following result:

Theorem 1.2. *Let Π be any LCL problem with locality $O(\log^* n)$ in the deterministic LOCAL model. Then Π can be also solved with locality $O(1)$ in the bounded-dependence model. Furthermore, the resulting finitely-dependent processes are invariant under subgraph isomorphism.*

Put otherwise, there is a finitely-dependent distribution over valid solutions of Π . Here the invariance under subgraph isomorphisms implies that, for any two graphs G, H that share some isomorphic subgraphs G' and H' such that their radius- $O(1)$ neighborhoods are still isomorphic, the finitely-dependent processes solving Π restricted to G' and H' are equal.

For any constant d , the task of coloring d -regular trees with $d + 1$ colors is a problem with locality $O(\log^* n)$ in the deterministic LOCAL model. Hence, we can **answer the open question** by Holroyd [40]:

Theorem 1.3. *For each $d \geq 2$, there is a finitely-dependent coloring with $d + 1$ colors in d -regular trees. Furthermore, the resulting process is invariant over subgraph isomorphisms.*

More specifically, there exists a finitely-dependent 4-coloring distribution of the infinite 3-regular tree that is invariant under automorphisms.

Theorem 1.2 also introduces a formal **barrier** for proving limitations on distributed quantum advantage. Recall that all current quantum-LOCAL lower bounds are, in essence, lower bounds in the non-signaling model. Before our work, there was a hope that we could discover a symmetry-breaking problem Π with the following properties: (1) its locality is $O(\log^* n)$ in deterministic LOCAL, and (2) we can show that its locality is $\Omega(\log^* n)$ in the non-signaling model, and therefore (3) Π cannot admit any distributed quantum advantage. However, our work shows that no such problem Π can exist. In particular, arguments related to non-signal distributions are not sufficient to exclude distributed quantum advantage in this region.

1.5 Locality in online and dynamic settings

Let us now switch gears and consider a very different line of work. Ghaffari et al. [35] introduced a *sequential* counterpart of the classical LOCAL model:

- **Deterministic SLOCAL model:** The nodes are processed in an adversarial order. When a node v is processed, the algorithm gets to see all information in its radius- $T(n)$ neighborhood. The algorithm has to label v with its local output, and the algorithm can also record other information in v , which it can exploit when other nodes near v are later processed.
- **Randomized SLOCAL model:** As above, but the algorithm has also access to a source of random bits.

Clearly SLOCAL is stronger than LOCAL. One key feature is that the processing order naturally breaks symmetry, and all symmetry-breaking LCLs can be solved with $O(1)$ locality in SLOCAL. One interpretation of our first contribution is that we also establish a new, unexpected similarity between SLOCAL and the bounded-dependence model: *both are able to solve any symmetry-breaking LCL with constant locality.*

A recent work [2] introduced the following models that capture the notion of locality also in the context of centralized dynamic graph algorithms and centralized online graph algorithms:

- **Deterministic dynamic-LOCAL model:** The adversary constructs the graph one edge at a time. The algorithm has a global view of the graph, but it has to maintain a feasible solution after each update. The algorithm is restricted so that after a modification at node v , it can only update the solution within distance $T(n)$ from v .
- **Deterministic online-LOCAL model:** The adversary presents the input graph one node at a time. When a node v is presented, the adversary also reveals the radius- $T(n)$ neighborhood of v . The algorithm has to then choose the output label of v .

Both SLOCAL and dynamic-LOCAL can be sandwiched between LOCAL and online-LOCAL [2]:

$$\begin{array}{ccc}
 \text{deterministic LOCAL} & \longrightarrow & \text{deterministic SLOCAL} \\
 \downarrow & & \downarrow \\
 \text{deterministic dynamic-LOCAL} & \longrightarrow & \text{deterministic online-LOCAL}
 \end{array} \tag{2}$$

There are also some problems in which deterministic online-LOCAL is much stronger than deterministic LOCAL: 3-coloring in bipartite graphs has locality $\tilde{O}(\sqrt{n})$ in deterministic LOCAL [16, 26] but $O(\log n)$ in online-LOCAL [2]; very recently Chang et al. [21] also showed that this is tight for online-LOCAL.

1.6 Contribution 2: connecting all models for LCLs in rooted trees

At first sight, the models discussed in Sections 1.3 and 1.5 seem to have very little in common; they seem to be orthogonal extensions of the classical deterministic LOCAL model. Furthermore, we have already seen evidence that online-LOCAL can be much stronger than deterministic LOCAL. Nevertheless, we can connect all these models in a unified manner, and prove strong limits on their expressive power. To this end, we introduce yet another model:

- **Randomized online-LOCAL model:** Like deterministic online-LOCAL, but the algorithm has also access to a source of random bits. Crucially, we are playing against an oblivious adversary (the adversary fixes the graph and the order of presenting nodes before the algorithm starts to flip coins).

Trivially, this is at least as strong as all models in diagram (2). However, the big surprise is that it is also at least as strong as all models in diagram (1). In Section 4 we prove:

Theorem 1.4. *Any LCL that can be solved in the non-signaling model with locality $T(n)$ can also be solved in the randomized online-LOCAL model with the same locality.*

Then we zoom into the case of rooted trees in Section 6 and prove:

Theorem 1.5. *Any LCL on rooted trees that can be solved in the randomized online-LOCAL model with locality $o(\log \log n)$ can also be solved in the deterministic LOCAL model with locality $O(\log^* n)$.*

Together with Theorem 1.2 and previously-known results, we also obtain the following corollaries:

Theorem 1.6. *In rooted trees, the following families of LCLs are the same:*

- *locality $O(\log^* n)$ in deterministic or randomized LOCAL or quantum-LOCAL,*
- *locality $O(1)$ in bounded-dependence model, non-signaling model, deterministic or randomized SLOCAL, dynamic-LOCAL, or deterministic or randomized online-LOCAL.*

In rooted trees, there is no LCL problem with locality between $\omega(\log^ n)$ and $o(\log \log n)$ in any of these models: deterministic and randomized LOCAL, quantum-LOCAL, bounded-dependence model, non-signaling model, deterministic and randomized SLOCAL, dynamic-LOCAL, and deterministic and randomized online-LOCAL.*

In particular, when we look at LCLs in rooted trees, $O(\log^* n)$ -round quantum algorithms are not any stronger than $O(\log^* n)$ -round classical algorithms. (However, it is still possible that there are some LCLs in trees that can be solved in $O(1)$ rounds in quantum-LOCAL and that require $\Theta(\log^* n)$ rounds in deterministic LOCAL; recall the discussion in Section 1.4.)

Theorem 1.5 can be seen as an extension of the result of Akbari et al. [2] that connects *deterministic* online-LOCAL with LOCAL for LCLs on rooted *regular* trees *without inputs*. Our result is applicable to the randomized online-LOCAL (and hence we can connect it with the non-signaling and quantum models), and it holds for any LCLs on rooted trees (possibly with irregularities and inputs). Recall that the presence of inputs makes a huge difference already in the case of directed paths [5, 22], and we need fundamentally new ideas, as we can no longer build on the classification from [9, 11].

1.7 Contribution 3: a new lower bound for coloring

So far we have connected randomized online-LOCAL with other models through simulation arguments that only work in rooted trees. Let us now put limitations on randomized online-LOCAL in a broader setting. Recall that in deterministic online-LOCAL we can 3-color bipartite graphs with locality $O(\log n)$ [2], and this is tight [21]. In Section 7 we show that randomness does not help:

Theorem 1.7. *3-coloring in bipartite graphs is not possible with locality $o(\log n)$ in the randomized online-LOCAL model.*

This demonstrates that even though randomized online-LOCAL is a very strong model, strong enough to simulate e.g. any non-signaling distribution, it is nevertheless possible to prove strong lower bounds in this model (which then imply lower bounds across the entire landscape of models).

1.8 Additional contributions

It is known that *deterministic* SLOCAL is strong enough to simulate *randomized* LOCAL. In Section 8 we show that the same holds also for dynamic-LOCAL:

Theorem 1.8. *Deterministic dynamic-LOCAL can simulate randomized LOCAL.*

The new model that we introduced, randomized online-LOCAL, is defined using an *oblivious* adversary. In Section 9 we show that this is also necessary: if we defined randomized online-LOCAL with an *adaptive* adversary, it would be as weak as the deterministic online-LOCAL model.

1.9 Discussion and open questions

By putting together all our contributions, the landscape shown in Fig. 1 emerges. We can sandwich all models between deterministic LOCAL and randomized online-LOCAL. Going downwards, we get symmetry breaking for free (as indicated by the blue arrows). And in the case of rooted trees, for the low-locality region $o(\log \log n)$ we can also navigate upwards (as indicated by the orange arrows). Table 1 gives some concrete examples of localities for LCL problems across the landscape of models.

Our work suggests a number of open questions; here are the most prominent ones:

Question 1. Is quantum-LOCAL stronger than randomized LOCAL for any LCL problem? In particular, is there any LCL problem with constant locality in quantum-LOCAL but super-constant locality in LOCAL? We conjecture that such a problem does not exist, but to show that, new proof techniques are needed.

Question 2. Can we *exclude* quantum advantage for any concrete LCL problem that has complexity $\Theta(\log^* n)$ in the classical LOCAL models? Again, we need new techniques to prove that, as non-signaling arguments cannot be used. Nevertheless, this could be easier to tackle than Question 1, as we can specifically engineer an LCL problem that would remain hard for quantum-LOCAL.

Question 3. Does shared global information or shared quantum state ever help with any LCL problem?

Question 4. Is it possible to simulate deterministic or randomized online-LOCAL in SLOCAL and LOCAL also in a broader graph class than rooted trees? If we could extend the result to *unrooted* trees, it would also imply new lower bounds for the widely-studied *sinkless orientation* problem [12, 15] across all models.

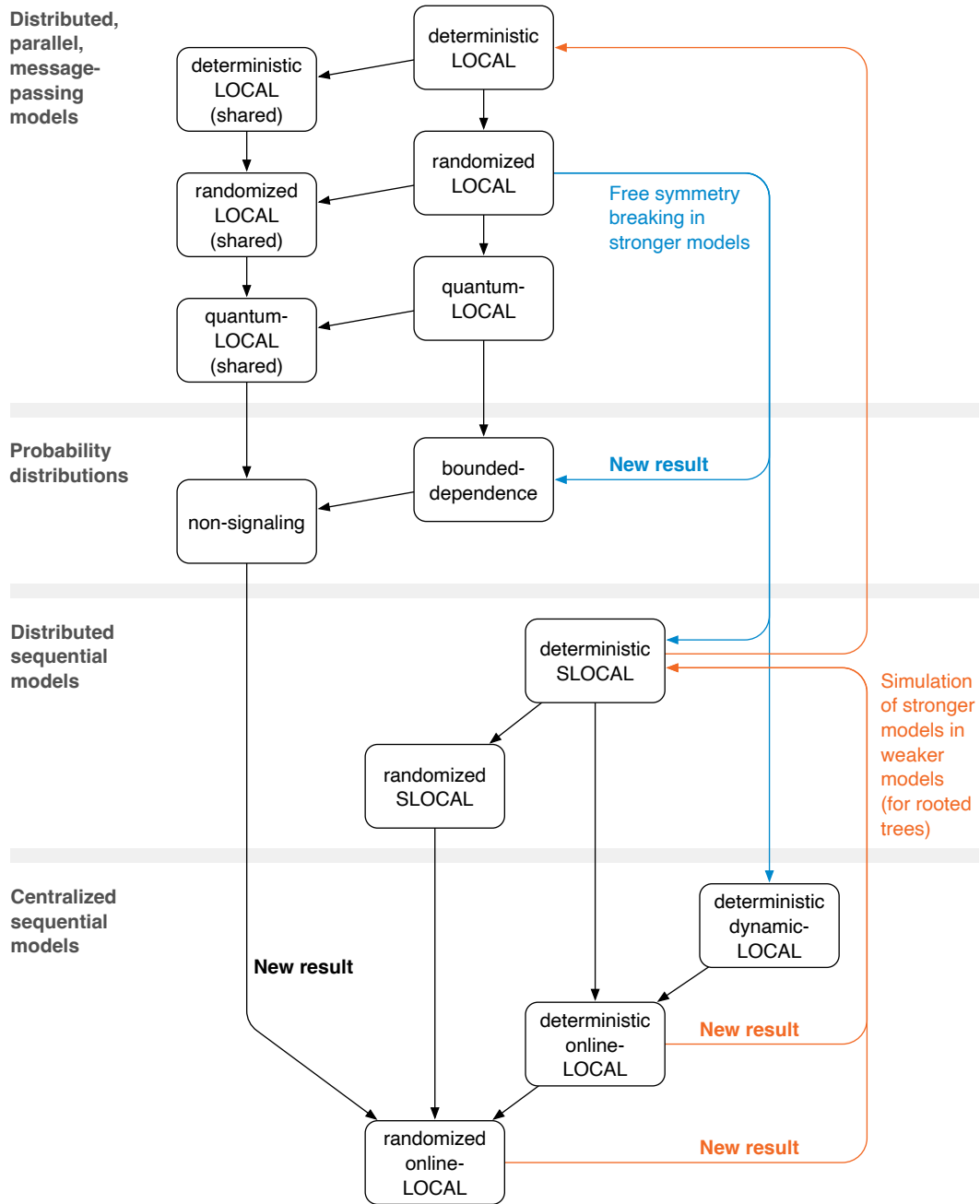


Figure 1: Landscape of models and their relations for LCL problems. A black arrow $X \rightarrow Y$ indicates that model Y is at least as strong as model X and can simulate any algorithm designed there. A blue arrow $X \rightarrow Y$ indicates that we additionally get symmetry-breaking for free: locality $O(\log^* n)$ in model X implies locality $O(1)$ in model Y . An orange arrow $X \rightarrow Y$ indicates that *at least in rooted trees* we can simulate highly-localized algorithms in model X using the weaker model Y so that e.g. locality $O(\log^* n)$ in model X implies locality $O(\log^* n)$ in model Y (see Theorem 6.1 and [35] for details).

Model	Symmetry-breaking problems		3-coloring in bipartite graphs	
Deterministic LOCAL	$\Theta(\log^* n)$	by definition	$\tilde{\Theta}(\sqrt{n})$	[26]
Randomized LOCAL	$\Theta(\log^* n)$	[18]	$\tilde{\Theta}(\sqrt{n})$	[26]
Quantum LOCAL	$O(\log^* n)$	trivial	$\tilde{\Theta}(\sqrt{n})$	[26]
Bounded-dependence	$O(1)$	Theorem 5.14	$\tilde{\Theta}(\sqrt{n})$	[26]
Deterministic SLOCAL	$O(1)$	[35]	$\tilde{O}(\sqrt{n}), n^{\Omega(1)}$	[2, 26]
Randomized SLOCAL	$O(1)$	[35]	$\tilde{O}(\sqrt{n}), n^{\Omega(1)}$	[2, 26]
Deterministic dynamic-LOCAL	$O(1)$	Theorem 8.2	$\tilde{O}(\sqrt{n}), \Omega(\log n)$	[21, 26]
Deterministic online-LOCAL	$O(1)$	[2]	$\Theta(\log n)$	[2, 21]
Randomized online-LOCAL	$O(1)$	[2]	$\Theta(\log n)$	Theorem 7.1

Table 1: Examples of localities across the models. Here *symmetry-breaking problems* refer to LCL problems with locality $\Theta(\log^* n)$ in the deterministic LOCAL model; this includes many classical problems such as maximal independent set, maximal matching, and $(\Delta + 1)$ -vertex coloring in graphs of maximum degree Δ .

2 Overview of techniques and key ideas

In this section, we give an overview of the techniques and key ideas that we use to prove our main results, and we also provide a roadmap to the rest of this paper. We note that our first contribution is presented in Section 5, while the second contribution comes before it in Section 4—the proofs are ordered this way, since Section 4 also develops definitions that will be useful in Section 5.

2.1 Bounded-dependence model can break symmetry (Section 5)

Let us first present an overview of the proofs of Theorems 1.2 and 1.3 from Section 1.4. We show that the bounded-dependence model can break symmetry with constant locality; that is, there is a finitely-dependent process for any symmetry-breaking LCL.

It is well known that any LCL problem Π that has complexity $O(\log^* n)$ in the LOCAL model has the following property: there exists a constant $k \in \mathbb{N}_+$ (that depends only on the hidden constant in $O(\log^* n)$) such that, if the graph is given a distance- k coloring (with sufficiently small number of colors) as an input, then Π is solvable in time $O(1)$ in the LOCAL model (using the distance- k coloring as a local assignment of identifiers) [19]. Furthermore, no knowledge of the size of the input graph is required.

Finding a distance- k coloring in SLOCAL trivially requires locality exactly k , implying the well known fact that LCLs of complexity $O(\log^* n)$ in the LOCAL model have complexity $O(1)$ in SLOCAL.

We prove that for each bounded-degree graph, there is a finitely-dependent process providing a distance- k coloring for constant k . Then, we can combine such a process with the LOCAL algorithm that solves the problem with locality $O(1)$ if a distance- k coloring is given, and we prove that the resulting process is still a finitely-dependent distribution. Furthermore, we also prove that all these processes are invariant under subgraph isomorphisms (even those that do not preserve node identifiers), meaning that, for any two graphs G and H sharing two isomorphic subgraphs with isomorphic radius- $O(1)$ neighborhoods, the restrictions of the finitely-dependent processes solving Π over G and H restricted to G' and H' are equal in law.

One of the key observations that we use is that LOCAL algorithms that do not exploit the specific assignment of node identifiers and do not depend on the size of the graph provide finitely-dependent distributions that are invariant under subgraph isomorphisms whenever the input labeling for the graphs is invariant under subgraph isomorphisms.

Overview. The cornerstone of our proof is a surprising result by Holroyd and Liggett [41] and its follow-up in [43], that state that there exist k -dependent distributions giving a q -coloring of the infinite path and of cycles for $(k, q) \in \{(1, 4), (2, 3)\}$ that are invariant under subgraph isomorphisms.

Recently, Holroyd has combined the finitely-dependent distributions of infinite paths to provide a finitely-dependent 4-coloring of the d -dimensional lattice [40]. Getting a translation invariant distribution is quite easy: First, use the distributions for the paths on each horizontal and vertical path obtaining a distance- k coloring (with k being a large enough constant) of the lattice with constantly many colors as shown in [41, Corollary 20]. Second, apply some LOCAL algorithm that starts from a distance- k coloring and reduces the number of colors to 4 while keeping the resulting distribution symmetric (e.g., the algorithms from [10, 16]). The major contribution of [40] is transforming such a distribution into a process that is invariant under subgraph isomorphisms. However, this *symmetrization* phase is quite specific to the considered topology.

We come up with a new approach that obtains similar results in all bounded-degree graphs through the following steps:

1. We show that the finitely-dependent coloring of paths and cycles can be combined to obtain finitely-dependent 3-coloring distributions of rooted pseudoforests of bounded-degree that are invariant under subgraph isomorphisms.
2. We observe that all graphs of bounded-degree admit a random decomposition in rooted pseudoforests that satisfies some required symmetry properties.
3. We prove that such random decomposition can be combined with the finitely-dependent 3-coloring of rooted pseudoforests to obtain finitely dependent distributions that give a $(\Delta + 1)$ -coloring of graphs of maximum degree Δ that are invariant under subgraph isomorphism.
4. We show that we can use this finitely dependent $(\Delta + 1)$ -coloring distributions to provide a distance- k coloring for graphs of maximum degree Δ which serves as an input for LOCAL algorithms solving any LCL Π of complexity $O(\log^* n)$ in LOCAL. Such combination results in finitely-dependent processes that are invariant under subgraph-isomorphisms and solve Π .

Notice that, in spirit, steps 1 to 3 are similar to the steps needed to produce a $(\Delta + 1)$ -coloring in time $O(\log^* n)$ in the LOCAL model [39, 55]: however, the detailed way these steps are obtained in the bounded-dependence model is quite different and requires a careful ad-hoc analysis.

Step 1: Finitely-dependent 3-coloring distributions of rooted pseudoforests. A pseudotree is a tree that contains at most one cycle. A rooted pseudotree is an oriented pseudotree in which each node has outdegree at most 1. A rooted pseudoforest is a collection of disjoint rooted pseudotrees. Let us now fix any rooted pseudoforest of maximum degree Δ . Consider the following process: each node v colors its indegree neighbors with a uniformly sampled permutation of the elements of $\{1, \dots, \text{indeg}(v)\}$. The graph G_i induced by nodes colored with color i is just a disjoint union of directed paths and cycles (see also Fig. 2 for an example) and, hence, admits a finitely-dependent 4-coloring given by [41] that is invariant under subgraph isomorphisms: if a node is isolated, it can deterministically join any of the G_i s, say G_1 . The sequence of graphs $(G_1, \dots, G_{\Delta_{\text{in}}})$ is said to be a random Δ_{in} -decomposition of the rooted pseudoforest. Furthermore, if two graphs G, H have isomorphic subgraphs G', H' (together with some constant-radius neighborhoods), the decompositions in directed paths and cycles induced in G' and H' have the same distribution (because node colors are locally chosen uniformly). We prove that the combination of the random decomposition and the finitely-dependent coloring yields a finitely-dependent 4Δ -coloring which is invariant under subgraph isomorphisms: by further combining such distribution with the Cole–Vishkin color reduction LOCAL algorithm [28, 39, 55] (that has complexity $O(\log^* k)$ with k being the size of the input coloring), we can obtain a finitely-dependent 3-coloring distribution for rooted pseudoforests of maximum degree Δ that is invariant under subgraph isomorphisms.

Steps 2–3: Finitely-dependent $(\Delta + 1)$ -coloring distribution of bounded-degree graphs. First, if the input graph is undirected, make it a directed graph by duplicating all edges and assigning both orientations to duplicated edges. Since a coloring of the nodes can be given in both cases equivalently, we focus on the directed case for simplicity. Second, consider the following process: each node v labels its outdegree edges with a uniformly sampled permutation of the elements of $\{1, \dots, \text{outdeg}(v)\}$. For each i , observe that edges labelled with i form a rooted pseudoforest (see also Fig. 3 for an example) that admits a finitely-dependent 3-coloring distribution (according to step 1) that is invariant under subgraph isomorphisms, inducing a random decomposition of the graph. Furthermore, if two graphs G, H have isomorphic subgraphs G', H' (together with some constant-radius neighborhoods), the decompositions induced in G' and H' have the same distribution (because edge labelings are locally chosen uniformly). We prove that the combination of the random decomposition and the finitely-dependent coloring yields a finitely-dependent 3^Δ -coloring of the bounded-degree graph which is invariant under subgraph isomorphisms: by further combining such distribution with a variant of the Cole–Vishkin color reduction LOCAL algorithm, we obtain a finitely-dependent $(\Delta + 1)$ -coloring distribution for bounded-degree graphs of maximum degree Δ that is invariant under subgraph isomorphisms.

Step 4: Finitely-dependent distribution solving Π . Consider any graph G of maximum degree Δ , and its k -th power graph defined as follows: simply add edges to G between each pair of nodes at distance at most k , where k is some large enough constant. Observe that G^k is a graph of maximum degree Δ^k . Now, step 3 implies that there is a finitely-dependent $(\Delta^k + 1)$ -coloring of G^k that is invariant under subgraph isomorphisms: such distribution yields a distance- k coloring of G . For any LCL Π that has complexity $O(1)$ in SLOCAL, we simply define a LOCAL algorithm \mathcal{A} solving Π that does not need node identifiers, does not depend on n , and has complexity $O(1)$: the combination of the input distance- k coloring of G with such an algorithm yields a finitely-dependent distribution solving Π that is invariant under subgraph isomorphisms. Basically, we take $k = T$ where T is the locality of some constant-time SLOCAL algorithm \mathcal{A} solving Π , and we define a LOCAL algorithm \mathcal{A}' as follows: nodes colored with color i perform T rounds of communication between round $(i - 1)T$ and round $iT - 1$ simulating the behavior of \mathcal{A} , while all other rounds of communications are useless. We prove that \mathcal{A}' is correct and provides the desired properties.

Random decomposition of a graph. In steps 1 and 3 we proceed in an analogous way: First, we construct a process that induces a random decomposition of a graph. Second, we consider finitely-dependent distributions of output labelings over the outputs of the random decomposition. The combination of the random decomposition and the finitely-dependent distributions gives rise to a process over the whole graph. In Section 5, we derive a general result (Lemma 5.8) which gives sufficient conditions on the random decomposition and the finitely-dependent distributions in order to ensure the final process is still finitely-dependent (possibly with symmetry properties). Lemma 5.8 is then the tool used in practice in steps 1 and 3.

Independent related work. Very recently, an independent and parallel work provided a finitely-dependent coloring of bounded-degree graphs with exponentially many colors (in the degree of the graph) [60]. The technique employed in [60] is very similar to ours: it exploits the decomposition of graphs in rooted pseudoforests, and then colors rooted pseudoforests using the finitely-dependent coloring of paths and cycles [41, 43]. However, [60] stops at the mere coloring problem and does not make use of color reduction algorithms, which are the key ingredient for extending results to all symmetry-breaking LCLs.

2.2 The non-signaling model can be simulated in randomized online-LOCAL (Section 4)

Let us now give the intuition behind the proof of Theorem 1.4 from Section 1.6: we show that the non-signaling model can be simulated in randomized online-LOCAL without any loss in the locality.

A randomized online-LOCAL algorithm is given in input the size of the input graph, and a distribution that is non-signaling beyond distance T and that solves some problem Π over some graph family \mathcal{F} . When the adversary picks any node v_1 and shows to the randomized online-LOCAL algorithm its radius- T neighborhood, the randomized online-LOCAL algorithm simply goes over all graphs of n nodes in \mathcal{F} until it finds one, say H_1 , that includes the radius- T neighborhood of v_1 : then, it samples an output according to the restriction of the non-signaling distribution in H_1 to v_1 . Notice that such distribution does not change if the topology of the graph changes outside the radius- T neighborhood of v_1 . Recursively, when the adversary picks the i -th node v_i , the randomized online-LOCAL algorithm goes over all graphs of n nodes in \mathcal{F} until it finds one, say H_i , that includes the union of radius- T neighborhoods of v_1, \dots, v_i (it must necessarily exist as the graph chosen by the adversary is a valid input): hence, it samples an output according to the restriction of the non-signaling distributions in H_i to v_i conditional on the outputs of v_1, \dots, v_{i-1} . We prove that the non-signaling property ensures that the algorithm described above fails with at most the same probability of failure of the non-signaling distribution.

2.3 Online-LOCAL can be simulated in SLOCAL for rooted trees (Section 6)

Next, we give an overview of the proof of Theorem 1.5 from Section 1.6: we show that a randomized online-LOCAL algorithm that solves an LCL problem in rooted trees with locality $o(\log^{(2)} n)$ can be simulated in the deterministic SLOCAL model with locality $O(1)$, and therefore also in the deterministic LOCAL model with locality $O(\log^* n)$.

Online-LOCAL algorithms can be seen as SLOCAL algorithms with a global memory. Notice that SLOCAL algorithms have instead an “incremental” memory, i.e., they keep track of an incremental sequence of intersecting neighborhoods $\mathcal{N}_1, \mathcal{N}_2, \dots$, where \mathcal{N}_i intersects \mathcal{N}_{i+1} .

The new ingredient we use in this section is *amnesiac online-LOCAL algorithms*: roughly speaking, an *amnesiac* algorithm is a deterministic online-LOCAL algorithm that, when processing a node v , uses information only coming from the connected component of the input graph that has been revealed so far to which v belongs, and nothing else. The steps to prove Theorem 1.5 are the following ones:

1. We first show that deterministic and randomized online-LOCAL algorithms that solve an LCL can be converted into *amnesiac* algorithms that solve the same LCL with some exponential overhead in the locality.
2. We then prove that, for an LCL on rooted trees, an SLOCAL algorithm can easily simulate the execution of *amnesiac* algorithms.

Step 1: Constructing amnesiac algorithms from deterministic online-LOCAL algorithms. To get some intuition on how we are going to do, start by considering an LCL problem Π on a family \mathcal{F} of graphs that is closed under disjoint graph union and node/edge removal, and a *deterministic* online-LOCAL algorithm \mathcal{A} solving Π with locality $T(n)$ in n -node graphs. Note that the output label \mathcal{A} chooses for a node may depend arbitrarily on the information the algorithm has seen so far. This is in contrast with *amnesiac* algorithms in which the output may depend only on the connected component of the graph that has been revealed so far to which the node belongs.

We now show how to turn algorithm \mathcal{A} into an algorithm whose output for *lonely* nodes depends only on the local topology and inputs, not on any previously-processed nodes; we call such algorithms 1-*amnesiac*. A node v is lonely if all nodes that belong to the radius T -neighborhood around v are new to algorithm \mathcal{A} , that is, \mathcal{A} does not know how v connects to other nodes it has seen before.

Let $N_1 = |\Sigma_{\text{out}}| |\Sigma_{\text{in}}|^n \cdot 2^{n^2} \cdot n^2$. Let \mathcal{G}_1 be the set of all possible subgraphs of any n -node graphs (from \mathcal{F}) with inputs (up to isomorphisms) that are the radius- $T(N_1)$ neighborhood of some node, which we call the *center* of the neighborhood. We denote $g_1 = |\mathcal{G}_1|$ and remark that $g_1 \leq 2^{n^2} |\Sigma_{\text{in}}|^n$. Now consider the following

experiment: Construct a simulation graph H_1 that consists of $|\Sigma_{\text{out}}| \cdot n$ copies of all graphs in \mathcal{G}_1 . The size of H_1 will be at most $N_1 = n^2 |\Sigma_{\text{out}}| \cdot g_1 \leq |\Sigma_{\text{out}}| |\Sigma_{\text{in}}|^n \cdot 2^{n^2} \cdot n^2$. Then reveal the center node of each of those neighborhood graphs to \mathcal{A} in an arbitrary order with locality $T(N_1)$. Now for each *type of radius- $T(N_1)$ neighborhood* \mathcal{T}_1 (i.e., any element of \mathcal{G}_1), there exists some output label $\sigma_{\mathcal{T}_1}$ that occurs at least n times; we call such neighborhoods *good* and such a label a *canonical labeling* for \mathcal{T}_1 . Afterwards, continue labelling nodes of H_1 executing \mathcal{A} following an arbitrary ordering of the nodes.

We can now imagine a new online-LOCAL algorithm \mathcal{B} that starts by running the above experiment with algorithm \mathcal{A} . Consider now an input graph G with n nodes, and let node v be revealed to \mathcal{B} along with its radius- $T'(n)$ neighborhood, with $T'(n) = T(N_1)$. Whenever v is a lonely node, algorithm \mathcal{B} finds an “unused” (that is, not previously chosen) good neighborhood in the experiment graph matching the radius- $T(N_1)$ neighborhood of v in G . It then takes that unused good neighborhood, identifies all nodes with the revealed input neighborhood, and then labels v accordingly. Since there are at least n good neighborhoods in the experiment graph matching the radius- $T(N_1)$ neighborhood of v in G , such unused good neighborhood must exist. In effect, algorithm \mathcal{B} cuts and pastes the neighborhood from the experiment graph to the actual input graph *without algorithm \mathcal{A} noticing*. For non-lonely nodes, \mathcal{B} just simulates what \mathcal{A} would do, following the adversarial order of the nodes presented to \mathcal{B} , which is always possible because the labels of lonely nodes come from valid online-LOCAL runs of \mathcal{A} .

The algorithm \mathcal{B} we just described labels the graph correctly by the virtue of \mathcal{A} working correctly. Moreover, when labeling a lonely node, it always labels it in the same way that depends only on the local structure and inputs of the graph: hence, \mathcal{B} is 1-amnesiac. We defer arguing about all formal details to Section 6. Clearly, there is an exponential overhead in the locality, as the locality of \mathcal{B} is $T'(n) = T(N_1) = T(2^{O(n^2)})$.

We now explain how to get a 2-amnesiac algorithm, that is, an algorithm that has always the same labels for the same types of connected components formed by the union of intersecting neighborhoods of two different nodes. To do so, we need to increase the size of the experiment graph. Let $N_2 = |\Sigma_{\text{out}}|^2 |\Sigma_{\text{in}}|^n \cdot 2^n \cdot n^2$, and redefine \mathcal{G}_1 with the radius- $T(N_2)$ neighborhoods. Instead of considering $|\Sigma_{\text{out}}| \cdot n$ many disjoint copies of elements of \mathcal{G}_1 , we now take $|\Sigma_{\text{out}}|^2 \cdot n$ copies. By what we argued before, there will be at least $n \cdot |\Sigma_{\text{out}}|$ good neighborhoods.

We now consider \mathcal{G}_2 to be the set of the union of two non-disjoint radius- $T(N_2)$ neighborhoods of two different nodes of any n -node graph (in \mathcal{F}), with all possible input labelings and orderings of the center nodes. Notice that the size of \mathcal{G}_2 is $g_2 \leq 2^{n^2} |\Sigma_{\text{in}}|^n$. We now take $|\Sigma_{\text{out}}| \cdot n$ many disjoint copies of all graphs in \mathcal{G}_2 , with the catch that the neighborhood of the center node that comes first in the processing order is chosen arbitrarily among the good neighborhoods. The resulting graph H_2 is the new experiment graph.

The size of the experiment graph H_2 is now $N_2 \leq |\Sigma_{\text{out}}| |\Sigma_{\text{in}}|^n \cdot 2^{n^2} \cdot n^3$. We ask \mathcal{A} to label first all nodes that are first in the adversarial order of each graph, and then all nodes that are second in the same order.

We now ask \mathcal{A} to label the second center node of each connected component. For all types of such connected components \mathcal{T}_2 there will be at least n identical labelings by the pigeonhole principle, which will be the canonical labelings $\sigma_{\mathcal{T}_2}$.

Now, the 2-amnesiac algorithm \mathcal{B} will start running the above experiment. When given in input an n -node graphs, it will output the canonical labeling $\sigma_{\mathcal{T}_1}$ for lonely nodes having their radius $T'(n) = T(N_2)$ neighborhood matching type \mathcal{T}_1 , and will output the canonical labeling $\sigma_{\mathcal{T}_2}$ for nodes seeing a connected component of type \mathcal{T}_2 when looking at their radius- $T'(n) = T(N_2)$ neighborhood. The correctness of this procedure follows analogously to that of 1-amnesiac algorithms. For other nodes, \mathcal{B} will just simulate \mathcal{A} with global memory.

The idea can be generalized up to n -amnesiac algorithms, which will not care about parts of the graph outside the current connected component. We defer to Lemma 6.4 all the formal details. The restriction to LCLs is necessary to prove correctness: if the amnesiac algorithm fails, it must fail locally, hence also the original online-LOCAL algorithm fails locally, reaching a contradiction.

How to deal with randomness. We now focus on \mathcal{A} being a randomized online-LOCAL algorithms. For deterministic online-LOCAL algorithms, we can adaptively pick the good neighborhoods before processing further. As the randomized online-LOCAL model calls for an oblivious adversary, we cannot do this.

We give the intuition on how to deal with this problem. The experiment graph H_n is *random*: it is constructed exactly as in the previous paragraph, with the catch that the good connected components have to be “guessed” uniformly at random each time. Clearly, the probability that the good guesses are enough is incredibly small: we would like to amplify it so that there is positive probability that the guesses are enough *and* the randomized online-LOCAL algorithm works correctly. The amplification is obtained by replicating k times the experiment graph H_n , guessing the good components for each copy of H_n independently. Since the randomness of the algorithm and the randomness of these guesses are independent for each copy of H_n , a simple union bound together with the inclusion/exclusion principle give positive probability that in at least one copy of H_n the guesses are correct and the algorithm works properly.

Hence, there exist an assignment of a random bit string to the randomized online-LOCAL algorithm, and a random bit string for “guessing” the good components in the k copies of the experiment graph, such that a fixed, deterministic realization of H_n and a deterministic online-LOCAL algorithm (that is given in input the random bit string) that works properly.

The n -amnesiac algorithm \mathcal{B} will just go over all possible definitions of H_n and of deterministic online-LOCAL algorithms (according to some order, e.g., lexicographically in the description) that meet our purpose until finding the good combination: there are only finitely many of them, given the size of the input graph.

The proof now reduces to the previous case. We defer to Lemma 6.5 all the formal details.

Step 2: From amnesiac algorithms to SLOCAL algorithms on rooted trees. In rooted trees, the fact that there is a consistent orientation of the edges towards the root can be heavily exploited. We can adapt results from [22, Section 7] to the SLOCAL model to show how to “cluster” in time $O(\alpha)$ a rooted tree in connected components (which are, in turn, rooted trees) so that the following properties are met:

1. All connected components have depth $\Theta(\alpha)$.
2. All leaves of a connected component that are not at distance $\Theta(\alpha)$ from the root of the component are real leaves of the original rooted tree.
3. All other leaves are either real leaves or roots of other connected components.

All formal details can be found in Lemma 6.7 and Section 6.

Suppose now we are given an n -amnesiac algorithm \mathcal{A} solving an LCL Π in rooted trees with locality $T(n)$ on n -node graphs. We now briefly describe how to construct an SLOCAL algorithm \mathcal{B} that solves Π with locality $O(T(n))$ in n -node rooted trees. We remind the reader that it is straightforward to prove that the composition of two SLOCAL algorithms with localities T_1 and T_2 yields an SLOCAL algorithm with locality $O(T_1 + T_2)$ [35].

We first consider an SLOCAL algorithm \mathcal{B}_1 that constructs a clustering with properties 1-3 above with $\alpha = O(T)$, where the hidden constant is large enough.

Then, we consider another SLOCAL algorithm \mathcal{B}_2 with locality $O(\alpha)$ that works as follows: each root of a connected component of the cluster precommits for a solution around the leaves of the component using \mathcal{A} : this can be done independently between different components if the localities are appropriately chosen. The only exception being the root of the original tree, which also outputs a solution for itself. Then, another SLOCAL algorithm \mathcal{B}_3 with locality $O(\alpha)$ will just output the precommitments on the right nodes.

The fact that we used \mathcal{A} on disjoint components of the graph, independently, and \mathcal{A} is a correct algorithm, ensures that a solution to the LCL exists and the “inner part” of all connected components can be filled with a correct solution: hence, \mathcal{B}_4 simply brute-forces a solution inside clusters with locality $O(\alpha)$. Since Π is an LCL, different components will have compatible solutions as \mathcal{B}_2 precommitted a solution around leaf nodes, which are roots of other connected components.

Then, \mathcal{B} is just the combination of $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$, has locality $O(\alpha) = O(T)$, and works correctly.

2.4 Lower bound on 3-coloring in randomized online-LOCAL (Section 7)

Let us now look at the proof of Theorem 1.7 from Section 1.7. We present a lower bound for 3-coloring in $\sqrt{n} \times \sqrt{n}$ grids, showing that it takes $\Omega(\log n)$ locality in the randomized online-LOCAL model.

Here it is important to assume that the adversary is oblivious, i.e., it cannot see the random decisions of the randomized algorithm. This lower bound complements a recent result of Chang et al. [21] showing the same bound for the deterministic online-LOCAL model.

In this proof, we use the notion of a b -value defined in [21] as a measure of the number of incompatible boundaries present in a region of a grid. We start with the assumption that a grid can be 3-colored with locality $o(\log n)$ and derive a contradiction. The high level idea is to construct two path segments below each other, where one path segment has a large count of incompatible boundaries (a high b -value) and the other segment has a low boundary count (incompatible b -value to the upper path). This forces an algorithm to make boundaries escape on the side between the two segments. We show that the boundary count is, however, too large compared to the distance between the two segments and thus the boundaries “cannot escape”.

Two difficulties arise in the randomized case compared to the deterministic lower bound: (i) In order to create a path with a large b -value we have to use a probabilistic construction that produces a segment with a large b -value with high probability; (ii) Since the first construction is probabilistic and the adversary oblivious, we cannot “see” the large b -value segment constructed in (i), that is, we can neither predict its position nor its size. We therefore need to use another probabilistic construction that positions the segment in a position that forces a contradiction (and which succeeds with constant probability).

2.5 Deterministic dynamic-LOCAL can derandomize randomized LOCAL (Section 8)

Let us now say a few words about the proof of Theorem 1.8 from Section 1.8: we show that deterministic dynamic-LOCAL can simulate randomized LOCAL.

The idea behind this result is the same as that for SLOCAL explained in Section 2.1. Any LCL problem Π that has complexity $O(\log^* n)$ in the LOCAL model has the following property: there exists a constant $k \in \mathbb{N}_+$ (that depends only on the hidden constant in $O(\log^* n)$) such that, if the input graph is given a distance- k coloring in the input, then Π is solvable in time $O(1)$ in the LOCAL model (using the distance- k coloring as a local assignment of identifiers) [19]. The proof consists in showing that dynamic-LOCAL can indeed provide a distance- k coloring in constant time.

2.6 Randomized online-LOCAL with adaptive adversary (Section 9)

As mentioned in Section 1.8, it is necessary to work with an oblivious adversary in randomized online-LOCAL: we show that an algorithm that works against an adaptive adversary can be turned into a deterministic algorithm.

We show that an adaptive adversary in randomized online-LOCAL is so strong that a succeeding randomized randomized online-LOCAL algorithm of locality T would admit a single random-bit string that outputs a good solution for all possible graphs of a given size: hence, a correct deterministic randomized online-LOCAL algorithm exists. Since deterministic online-LOCAL algorithms of locality T for graphs of n nodes are only finitely many, the deterministic online-LOCAL algorithm, in the initialization phase, can go over all of them until it finds the one working for all graphs of n nodes: it then uses that one.

2.7 Randomized online-LOCAL in paths and cycles (Section 10)

Our final technical part shows that LCL problems in paths and cycles have complexity either $O(1)$ or $\Theta(n)$ in the randomized online-LOCAL model; moreover the locality is $O(1)$ in randomized online-LOCAL if and only if it is $O(\log^* n)$ in the deterministic LOCAL model. Together with prior work, this also shows that locality of an LCL problem in paths and cycles is *decidable* across all models [5] (with the caveat that we cannot distinguish between $O(1)$ and $\Theta(\log^* n)$ for quantum-LOCAL).

The proof is a reworking of its deterministic variant from [2]. The main take-home message of this result is the following: cycles are not a fundamental obstacle for simulating randomized online-LOCAL in weaker models. Hence, there is hope for generalizing the simulation result of Section 6 from trees to a broader class of graphs.

2.8 Quantum models vs. non-signaling and bounded-dependence models (Appendix A)

Appendix A aims at serving a dual purpose. First, it aims at formally introducing the non-signaling model based on the non-signaling principle, and at explaining *why* it is more powerful than the quantum-LOCAL model. This is not a new result, but included for completeness and to clarify the early works of Gavaille et al. [33] and Arfaoui and Fraigniaud [3].

Second, it formally introduces the bounded-dependence model based on finitely-dependent processes and argues why the relations in diagram (1) hold, and in particular why quantum-LOCAL without shared quantum state is contained not only in the non-signaling model but also in the bounded-dependence model. While all the ingredients are well-known, to our knowledge this relation between the quantum-LOCAL model and the bounded-dependence model is not made explicit in the literature before.

3 Preliminaries

We use the notation $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. For any positive integer $n \in \mathbb{N}_+$, we denote the set $\{1, \dots, n\}$ by $[n]$.

Graphs. In this work, a graph $G = (V, E)$ can be either directed ($E \subseteq V^2$) or undirected ($E \subseteq \binom{V}{2}$). If the set of nodes and the set of edges are not specified, we refer to them by $V(G)$ and $E(G)$, respectively. For any edge $e = (u, v) \in E(G)$ we say that e is directed from u to v and is incident to u and to v (the latter holds also for undirected edges). All graphs in this paper are simple graphs without self-loops unless differently specified. The degree of a node v is the number of edges that are incident to v and is denoted by $\deg_G(v)$, or simply by $\deg(v)$ when G is clear from the context. The indegree of a node v is the number of directed edges that are directed towards v and is denoted by $\text{indeg}_G(v)$, while the outdegree of v is the number of directed edges that are incident to v but directed to some other vertex and is denoted by $\text{outdeg}_G(v)$. Again, we omit the suffix G if the graph is clear from the context.

If G is a subgraph of H , we write $G \subseteq H$. For any subset of nodes $A \subseteq V$, we denote by $G[A]$ the subgraph induced by the nodes in A . For any nodes $u, v \in V$, $\text{dist}_G(u, v)$ denotes the distance between u and v in G (i.e., the number of edges of any shortest path between u and v in G —it doesn't need to be a directed path); if u and v are disconnected, then $\text{dist}_G(u, v) = +\infty$. If G is clear from the context, we may also simply write $\text{dist}(u, v) = \text{dist}_G(u, v)$. Also, for subset of nodes $A, B \subseteq V$ and any node $v \in V$, we can define $\text{dist}_G(v, A) = \min_{u \in A} \text{dist}_G(u, v)$ and, by extension, $\text{dist}_G(A, B) = \min_{u \in A} \text{dist}_G(u, B)$. We assume that $\text{dist}_G(A, \emptyset) = +\infty$. Similarly, for subgraphs $G_1, G_2 \subseteq G$, we define $\text{dist}_G(G_1, G_2) = \text{dist}_G(V(G_1), V(G_2))$: here, we also assume $\text{dist}_G(G_1, \emptyset) = +\infty$ where \emptyset is now the empty graph. For $T \in [n]$, the T -neighborhood of a node $u \in V$ of a graph G is the set $\mathcal{N}_T(u, G) = \{v \in V \mid \text{dist}_G(u, v) \leq T\}$. The T -neighborhood of a subset $A \subseteq V$ is the set $\mathcal{N}_T(A, G) = \{v \in V \mid \exists u \in A : \text{dist}_G(u, v) \leq T\}$. Similarly, the T -neighborhood of a subgraph

$H \subseteq G$ is the set $\mathcal{N}_T(H, G) = \{v \in V \mid \exists u \in V(H) : \text{dist}_G(u, v) \leq T\}$. If G is clear from the context, we just write $\mathcal{N}_T(u)$, $\mathcal{N}_T(A)$, and $\mathcal{N}_T(H)$. If $u \notin V$, $A \cap V = \emptyset$, or $V(H) \cap V(G) = \emptyset$, then the neighborhood $\mathcal{N}_T(u) = \emptyset$, $\mathcal{N}_T(A) = \emptyset$, or $\mathcal{N}_T(H) = \emptyset$, respectively. We make use of some graph operations: For any two graphs G, H , we denote by $G \cap H$ the intersection graph defined by $G \cap H = (V(G) \cap V(H), E(G) \cap E(H))$. The graph union is defined by $G \cup H = (V(G) \cup V(H), E(G) \cup E(H))$. Moreover, the graph difference is the graph $G \setminus H = (V(G) \setminus V(H), E(G) \setminus E(H))$.

Finally, for any two graphs G and H , we write $G \sim_f H$ to denote that G and H are isomorphic and $f : V(G) \rightarrow V(H)$ is an isomorphism.

Labeling problems. We start with the notion of labeling problem.

Definition 3.1 (Labeling problem). Let Σ_{in} and Σ_{out} two sets of input and output labels, respectively. A *labeling problem* Π is a mapping $(G, \lambda_{\text{in}}) \mapsto \{\lambda_{(\text{out}, i)}\}_{i \in I}$, with I being a discrete set of indexes, that assigns to every graph G with any input labeling $\lambda_{\text{in}} : V(G) \rightarrow \Sigma_{\text{in}}$ a set of permissible output vectors $\lambda_{(\text{out}, i)} : V(G) \rightarrow \Sigma_{\text{out}}$ that might depend on (G, λ_{in}) . The mapping must be closed under graph isomorphism, i.e., if $\varphi : V(G) \rightarrow V(G')$ is an isomorphism between G and G' , and $\lambda_{(\text{out}, i)} \in \Pi((G', \lambda_{\text{in}}))$, then $\lambda_{(\text{out}, i)} \circ \varphi \in \Pi((G, \lambda_{\text{in}} \circ \varphi))$.

A labeling problem can be thought as defined for *any* input graph of *any* number of nodes. If the set of permissible output vectors is empty for some input (G, λ_{in}) , we say that the problem is not solvable on the input (G, λ_{in}) : accordingly, the problem is solvable on the input (G, λ_{in}) if $\Pi(G, \lambda_{\text{in}}) \neq \emptyset$.

One observation on the generality of definition of labeling problem follows: one can actually consider problems that require to output labels on edges.

We actually focus on labeling problems where, for any input graph, an output vector λ_{out} is permissible if and only if the restrictions of the problem on any local neighborhoods can be solved and there exist compatible local permissible output vectors whose combination provides λ_{out} . This concept is grasped by the notion of locally checkable labeling (LCL) problems, first introduced by Naor and Stockmeyer [52]. For any function $f : A \rightarrow B$ and any subset $A' \subseteq A$, let us denote the restriction of f to A' by $f \upharpoonright_{A'}$. Furthermore, we define a centered graph to be a pair (H, v_H) where H is a graph and $v_H \in V(H)$ is a vertex of H that we name the *center* of H . The *radius* of a centered graph is the maximum distance from v_H to any other node in H .

Definition 3.2 (Locally checkable labeling problem). Let $r, \Delta \in \mathbb{N}$. Let Σ_{in} and Σ_{out} two finite sets of input and output labels, respectively, and Π a labeling problem. Π is *locally checkable* with checking radius r if there exists a family $\mathcal{S} = \{(H, v_H), \bar{\lambda}_{\text{in}}, \bar{\lambda}_{\text{out}}\}_{i \in I}$ of tuples, where (H, v_H) is a centered graph of radius at most r and maximum degree at most Δ , $\bar{\lambda}_{\text{in}} : V(H) \rightarrow \Sigma_{\text{in}}$ is an input labeling for H , $\bar{\lambda}_{\text{out}} : V(H) \rightarrow \Sigma_{\text{out}}$ is an output labeling for H (which can depend on $\bar{\lambda}_{\text{in}}$) with the following property

- for any input (G, λ_{in}) to Π with $\text{deg}(G) \leq \Delta$, an output vector $\lambda_{\text{out}} : V(G) \rightarrow \Sigma_{\text{out}}$ is permissible (i.e., $\lambda_{\text{out}} \in \Pi((G, \lambda_{\text{in}}))$) if and only if, for each node $v \in V(G)$, the tuple $((G[\mathcal{N}_r(v)]), \lambda_{\text{in}} \upharpoonright_{\mathcal{N}_r(v)}, \lambda_{\text{out}} \upharpoonright_{\mathcal{N}_r(v)})$ belongs to \mathcal{S} (up to graph isomorphisms).

Notice that the family \mathcal{S} can be always thought to be finite up to graph isomorphisms, as r and Δ are fixed and the set of input/output labels are finite. We now define the computational models we work in.

The port-numbering model. A port-numbered network is a triple $N = (V, P, p)$ where V is the set of nodes, P is the set of *ports*, and $p : P \rightarrow P$ is a function specifying connections between ports. Each element $x \in P$ is a pair (v, i) where $v \in V$, $i \in \mathbb{N}_+$. The connection function p between ports is an involution, that is, $p(p(x)) = x$ for all $x \in P$. If $(v, i) \in P$, we say that (v, i) is port number i in node v . The degree of a node v in the network N is $\text{deg}_N(v)$ is the number of ports in v , that is, $\text{deg}_N(v) = |\{i \in \mathbb{N} : (v, i) \in P\}|$. Unless

otherwise mentioned, we assume that port numbers are consecutive, i.e., the ports of any node $v \in V$ are $(v, 1), \dots, (v, \deg_N(v))$. Clearly, a port-numbered network identifies an *underlying graph* $G = (V, E)$ where, for any two nodes $u, v \in V$, $\{u, v\} \in E$ if and only if there exists ports $x_u, x_v \in P$ such that $p(x_u) = x_v$. Clearly, the degree of a node $\deg_N(v)$ corresponds to $\deg_G(v)$.

In the port-numbering model we are given distributed system consisting of a port-numbered network of $|V| = n$ processors (or nodes) that operates in a sequence of synchronous rounds. In each round the processors may perform unbounded computations on their respective local state variables and subsequently exchange of messages of arbitrary size along the links given by the underlying input graph. Nodes identify their neighbors by using ports as defined before, where port assignment may be done adversarially. Barring their degree, all nodes are identical and operate according to the same local computation procedures. Initially all local state variables have the same value for all processors; the sole exception is a distinguished local variable $x(v)$ of each processor v that encodes input data.

Let Σ_{in} be a set of input labels. The input of a problem is defined in the form of a labeled graph (G, x) where $G = (V, E)$ is the system graph, V is the set of processors (hence it is specified as part of the input), and $x: V \rightarrow \Sigma_{\text{in}}$ is an assignment of an input label $\lambda_{\text{in}}(v) \in \Sigma_{\text{in}}$ to each processor v . The output of the algorithm is given in the form of a vector of local output labels $\lambda_{\text{out}}: V \rightarrow \Sigma_{\text{out}}$, and the algorithm is assumed to terminate once all labels $\lambda_{\text{out}}(v)$ are definitely fixed. We assume that nodes and their links are fault-free. The local computation procedures may be randomized by giving each processor access to its own set of random variables; in this case, we are in the *randomized* port-numbering model as opposed to the *deterministic* port-numbering model.

The running time of an algorithm is the number of synchronous rounds required by all nodes to produce output labels. If an algorithm running time is T , we also say that the algorithm has locality T . Notice that T can be a function of the size of the input graph. We say that a problem Π over some graph family \mathcal{F} has complexity T in the port-numbering model if there is a port-numbering algorithm running in time T that solves Π over \mathcal{F} . If the algorithm is randomized, we also require that the failure probability is at most $1/n$, where n is the size of the input graph.

We remark that the notion of an (LCL) problem is a graph problem, and does not depend on the specific model of computation we consider (hence, the problem cannot depend on, e.g., port numbers).

The LOCAL model. The LOCAL model was first introduced by Linial [49]: it is just the port-numbering model augmented with an assignment of unique identifiers to nodes. Let $c \geq 1$ be a constant, and let Σ_{in} be a set of input labels. The input of a problem is defined in the form of a labeled graph (G, x) where $G = (V, E)$ is the system graph, V is the set of processors (hence it is specified as part of the input), and $x: V \rightarrow [n^c] \times \Sigma_{\text{in}}$ is an assignment of a *unique* identifier $\text{id}(v) \in [n^c]$ and of an input label $\lambda_{\text{in}}(v) \in \Sigma_{\text{in}}$ to each processor v . The output of the algorithm is given in the form of a vector of local output labels $\lambda_{\text{out}}: V \rightarrow \Sigma_{\text{out}}$, and the algorithm is assumed to terminate once all labels $\lambda_{\text{out}}(v)$ are definitely fixed. We assume that nodes and their links are fault-free. The local computation procedures may be randomized by giving each processor access to its own set of random variables; in this case, we are in the *randomized* LOCAL (randomized LOCAL) model as opposed to *deterministic* LOCAL (deterministic LOCAL). Notice that the knowledge of n makes the randomized port-numbering model roughly equivalent to the randomized LOCAL model, as unique identifiers can be produced with high probability. We say that a problem Π over some graph family \mathcal{F} has complexity T in the LOCAL model if there is a LOCAL algorithm running in time T that solves Π over \mathcal{F} . If the algorithm is randomized, we also require that the failure probability is at most $1/n$, where n is the size of the input graph.

The sequential LOCAL model. The sequential LOCAL model was first introduced by [35]: it is a sequential version of the LOCAL model. Nodes are processed according to an adversarial order $\sigma = v_1, \dots, v_n$. When

processing a node v_i , a T -round algorithm collects all inputs in the radius- T neighborhood of v_i (including outputs of those $v_j \in \mathcal{N}_T(V_i)$ for $j < i$): we say that such an algorithm has complexity T . Note that the algorithm might store all inputs in $\mathcal{N}_T(v_i)$ in the output of v_i ; hence, when processing v_i , it can see the input of v_j , $j < i$, if and only if there is a subsequence of nodes $\{v_{h_k}\}_{k \in [m]}$ with $j = h_k < h_{k+1} < \dots < h_{k_m} = i$ such that $v_{h_k} \in \mathcal{N}_T(v_{h_{k+1}})$ for all $k \in [m]$.

If the algorithm is given an infinite random bit string, we talk about the randomized SLOCAL model, as opposed to the deterministic SLOCAL model. We assume that the adversarial order according to which nodes are processed is oblivious to the random bit string, as in the original definition of the model. We say that a problem Π over some graph family \mathcal{F} has complexity T in the SLOCAL model if there is an SLOCAL algorithm running in time T that solves Π over \mathcal{F} . If the algorithm is randomized, we also require that the failure probability is at most $1/n$, where n is the size of the input graph.

The online-LOCAL model. The (deterministic) online-LOCAL model was introduced in [2]. This is a centralized model where the algorithm initially knows only the set of nodes of the input graph G . The nodes are processed with respect to an adversarial input sequence $\sigma = v_1, v_2, \dots, v_n$, such that the label of the v_i depends on $G_i = G[\bigcup_{j=1}^i \mathcal{N}_T(v_j)]$, i.e., the subgraph induced by the radius- T neighborhoods of v_1, v_2, \dots, v_i (including all input data).

We define the randomized online-LOCAL model as a randomized variant of the online-LOCAL model where the label assigned by the algorithm to v_i is a random outcome. Note that this model is oblivious to the randomness used by the algorithm. In particular this means that the graph $G \setminus G_i$ cannot be changed depending on the label assigned to v_i . One could also define the randomized online-LOCAL model in an adaptive manner, but it turns out that this is equivalent to the deterministic online-LOCAL model as we show in Section 9. We say that a problem Π over some graph family \mathcal{F} has complexity T in the online-LOCAL (randomized online-LOCAL) model if there is an online-LOCAL (randomized online-LOCAL) algorithm running in time T that solves Π over \mathcal{F} . If the algorithm is randomized, we also require that the failure probability is at most $1/n$, where n is the size of the input graph.

4 Simulation of the non-signaling model in randomized online-LOCAL

4.1 Framework

In this section we give the necessary framework to define the non-signaling model and is largely inspired by the definitions given in [26, 33]. Next definition introduces the concept of outcome.

Definition 4.1 (Outcome). Let Σ_{in} and Σ_{out} be two sets of input and output labels, respectively, and let \mathcal{F} be a family of graphs. An *outcome* \mathcal{O} over \mathcal{F} is a mapping $(G, x) \mapsto \{(\lambda_{(\text{out},i)}, p_i)\}_{i \in I}$, with I being a discrete set of indexes, assigning to every input graph $G \in \mathcal{F}$ with any input data $x = (\text{id} : V(G) \rightarrow [|V(G)|]^c, \lambda_{\text{in}} : V(G) \rightarrow \Sigma_{\text{in}})$, a discrete probability distribution $\{p_i\}_{i \in I}$ over output vectors $\lambda_{(\text{out},i)} : V(G) \rightarrow \Sigma_{\text{out}}$ such that:

1. for all $i \in I$, $p_i > 0$;
2. $\sum_{i \in I} p_i = 1$;
3. p_i represents the probability of obtaining $\lambda_{(\text{out},i)}$ as the output vector of the distributed system.

We say that an outcome \mathcal{O} over some graph family \mathcal{F} *solves* problem Π over \mathcal{F} *with probability* p if, for every $G \in \mathcal{F}$ and any input data $x = (\text{id}, \lambda_{\text{in}})$, it holds that

$$\sum_{\substack{(\lambda_{(\text{out},i)}, p_i) \in \mathcal{O}((G, x)) : \\ \lambda_{(\text{out},i)} \in \Pi((G, \lambda_{\text{in}}))}} p_i \geq p.$$

When $p = 1$, we will just say that \mathcal{O} solves problem Π over the graph family \mathcal{F} .

The next computational model tries to capture the fundamental properties of any *physical* computational model (in which one can run either deterministic, random, or quantum algorithms) that respects causality. The defining property of such a model is that, for any two (labeled) graphs (G_1, x_1) and (G_2, x_2) that share some identical subgraph (H, y) , every node u in H must exhibit identical behavior in G_1 and G_2 as long as its *local view*, that is, the set of nodes up to distance T away from u together with input data and port numbering, is fully contained in H . As the port numbering can be computed with one round of communication through a fixed procedure (e.g., assigning port numbers $1, 2, \dots, \deg(v)$ based on neighbor identifiers in ascending order) and we care about asymptotic bounds, we will omit port numbering from the definition of local view.

The model we consider has been introduced by [33]. In order to proceed, we first define the *non-signaling* property of an outcome. Let $T \geq 0$ be an integer, and I a set of indices. For any set of nodes V , subset $S \subseteq V$, and for any input $(G = (V, E), x)$, we define its *T-local view* as the set

$$\mathbf{v}_T(G, x, S) = \{(u, x(u)) \mid \exists u \in V, v \in S \text{ such that } \text{dist}_G(u, v) \leq T\},$$

where $\text{dist}_G(u, v)$ is the distance in G . Furthermore, for any subset of nodes $S \subseteq V$ and any output distribution $\{\lambda_{(\text{out}, i)}, p_i\}_{i \in I}$, we define the *marginal distribution* of $\{\lambda_{(\text{out}, i)}, p_i\}_{i \in I}$ on set S as the unique output distribution defined as follows: for any $\lambda_{\text{out}} : V(G) \rightarrow \Sigma_{\text{out}}$, the probability of $\lambda_{\text{out}} \upharpoonright_S$ on S is given by

$$p(\lambda_{\text{out}}, S) = \sum_{i \in I : \lambda_{\text{out}} \upharpoonright_S = \lambda_{(\text{out}, i)} \upharpoonright_S} p_i,$$

where $\lambda_{\text{out}} \upharpoonright_S$ and $\lambda_{(\text{out}, i)} \upharpoonright_S$ are the restrictions of λ_{out} and $\lambda_{(\text{out}, i)}$ on S , respectively.

Definition 4.2 (Non-signaling outcome). Let \mathcal{F} be a family of graphs. An outcome $\mathcal{O} : (G, x) \mapsto \{(\lambda_{(\text{out}, i)}, p_i)\}_{i \in I}$ over \mathcal{F} is *non-signaling* beyond distance $T = T(G, x)$ if for any pair of inputs $(G_1 = (V_1, E_1), x_1)$, $(G_2 = (V_2, E_2), x_2)$, with the same number of nodes, such that $\mathbf{v}_{T(G_1, x_1)}(G_1, x_1, S)$ is isomorphic to $\mathbf{v}_{T(G_2, x_2)}(G_2, x_2, S)$ and $G_1, G_2 \in \mathcal{F}$, the output distributions corresponding to these inputs have identical marginal distributions on the set S .

Definition 4.2 is also the more general definition for the locality of an outcome: an outcome \mathcal{O} has locality T if it is non-signaling beyond distance T .

The φ -LOCAL model. The φ -LOCAL model is a computational model that produces non-signaling outcomes over some family of graphs \mathcal{F} . Let $p \in [0, 1]$. A problem Π over some graph family \mathcal{F} has complexity T (and success probability p) if there exists an outcome \mathcal{O} that is non-signaling beyond distance T which solves Π over \mathcal{F} (with probability at least p).

As every (deterministic or randomized) algorithm running in time at most T in the LOCAL model produces an outcome which has locality T , we can provide lower bounds for the LOCAL model by proving them in the φ -LOCAL model.

Notice that algorithms in the LOCAL model can be always thought as producing outputs for *any* input graph: when the computation at any round is not defined for some node, we can make the node output some garbage label, say \perp . If we require that also outcomes are defined for every possible graph, then we are restricting the power of φ -LOCAL because outcomes must be defined accordingly. This gives rise to a slightly weaker model, the non-signaling model, which was considered in other works such as [26] and is still stronger than any classical or quantum variation of the LOCAL model.

Theorem 4.3. Let Π be any LCL problem over any family of graphs \mathcal{F} . Let $\mathcal{O} : (G, \lambda_{\text{in}}) \mapsto \{(\lambda_{(\text{out}, h)}, p_i)\}_{h \in H}$ be an outcome over \mathcal{F} which solves Π with failure probability at most ε and is non-signaling at distance greater than T . There exists a randomized online-LOCAL algorithm \mathcal{A} with complexity T that solves Π over \mathcal{F} and has failure probability at most ε .

We want to design a randomized online-LOCAL algorithm \mathcal{A} with complexity T that solves Π over \mathcal{F} and has failure probability at most ε that somehow simulates \mathcal{O} . We need to assume that the number of nodes of the input graph is known by \mathcal{A} .

Fix any graph $G \in \mathcal{F}$ of n nodes and any input labeling λ_{in} , and fix any sequence of nodes v_1, \dots, v_n the adversary might choose to reveal to the algorithm.

The adversary initially shows to the algorithm $G[\mathcal{N}_T(v_1)]$ (including input labels and identifiers) and asks to label v_1 . For the algorithm, it is sufficient to find any graph $H_1 \in \mathcal{F}$ of n nodes that contains a subgraph isomorphic to $G[\mathcal{N}_T(v_1)]$ (including input labels and identifiers): the non-signaling property ensures that the restriction of the output distribution $\mathcal{O}(H_1, \lambda_{\text{in}})$ over $\mathcal{N}_T(v_1)$ does not change if the topology of the graph outside $G[\mathcal{N}_T(v_1)]$ differs. Notice that H_1 exists necessarily as G itself is such a graph.

For any $\lambda_{\text{out}} : V(G) \rightarrow \Sigma_{\text{out}}$, the probability that \mathcal{A} labels v_1 with $\lambda_{\text{out}}(v_1)$ is

$$p(\lambda_{\text{out}}, v_1) = \sum_{k: \lambda_{\text{out}} \upharpoonright_{\{v_1\}} = \lambda_{(\text{out},k)} \upharpoonright_{\{v_1\}}} p_k,$$

where $\{(\lambda_{(\text{out},k)}, p_k)\}_{k \in K_1} = \mathcal{O}(H_1, \lambda_{\text{in}})$.

Let Λ_i be the random variable yielding the label assigned to v_i by \mathcal{A} . In general, for any $\lambda_{\text{out}} : V(G) \rightarrow \Sigma_{\text{out}}$ let $p(\lambda_{\text{out}}, v_i)$ denote the probability that $\Lambda_i = \lambda_{\text{out}}(v_i)$. Assume now that $G[\mathcal{N}_T(v_{i+1})]$ is shown to the algorithm. Conditional on $\Lambda_1, \dots, \Lambda_i$, for any $\lambda_{\text{out}} : V(G) \rightarrow \Sigma_{\text{out}}$ such that $\lambda_{\text{out}}(v_j) = \Lambda_j$ for all $j = 1, \dots, i$, the probability that $\Lambda_{i+1} = \lambda_{\text{out}}(v_{i+1})$ is

$$p(\lambda_{\text{out}}, v_{i+1}) = \sum_{\substack{k: \lambda_{\text{out}} \upharpoonright_{\{v_j\}} = \lambda_{(\text{out},k)} \upharpoonright_{\{v_j\}} \\ \forall 1 \leq j \leq i+1}} \frac{p_k}{\prod_{1 \leq l \leq i} p(\lambda_{\text{out}}, v_l)},$$

where $\{(\lambda_{(\text{out},k)}, p_k)\}_{k \in K_{i+1}} = \mathcal{O}(H_{i+1}, \lambda_{\text{in}})$ for any graph $H_{i+1} \in \mathcal{F}$ of n nodes that contains a subgraph isomorphic to $G[\cup_{i=1}^{i+1} \mathcal{N}_T(v_j)]$ (including input labels and identifiers). The specific choice of H_{i+1} does not matter since the property of a non-signaling outcome is exactly that the probability distribution of any output over any subset S of nodes does not change if the topology of the graph is modified “far enough” from S : here, $S = \{v_1, \dots, v_{i+1}\}$.

Now, consider $\mathcal{O}(G, \lambda_{\text{in}}) = \{(\lambda_{(\text{out},h)}, p_h)\}_{h \in H}$ for the right input graph G , and take any

$$(\lambda_{(\text{out},h^*)}, p_{h^*}) \in \mathcal{O}(G, \lambda_{\text{in}}).$$

The probability that the outcomes of $\Lambda_1, \dots, \Lambda_n$ give exactly $\lambda_{(\text{out},h^*)}$ is

$$\begin{aligned} & \Pr \left[\Lambda_1 = \lambda_{(\text{out},h^*)}(v_1), \dots, \Lambda_n = \lambda_{(\text{out},h^*)}(v_n) \right] \\ &= \Pr \left[\Lambda_n = \lambda_{(\text{out},h^*)}(v_n) \mid \Lambda_1 = \lambda_{(\text{out},h^*)}(v_1), \dots, \Lambda_{n-1} = \lambda_{(\text{out},h^*)}(v_{n-1}) \right] \\ & \quad \cdot \Pr \left[\Lambda_{n-1} = \lambda_{(\text{out},h^*)}(v_{n-1}) \mid \Lambda_1 = \lambda_{(\text{out},h^*)}(v_1), \dots, \Lambda_{n-2} = \lambda_{(\text{out},h^*)}(v_{n-2}) \right] \\ & \quad \dots \\ & \quad \cdot \Pr \left[\Lambda_1 = \lambda_{(\text{out},h^*)}(v_1) \right] \\ &= p(\lambda_{(\text{out},h^*)}, v_1) \cdot \dots \cdot p(\lambda_{(\text{out},h^*)}, v_n) \\ &= \prod_{1 \leq l \leq n-1} p(\lambda_{\text{out}}, v_l) \cdot \sum_{\substack{k_n: \lambda_{(\text{out},h^*)} \upharpoonright_{\{v_j\}} = \lambda_{(\text{out},k_j)} \upharpoonright_{\{v_j\}} \\ \forall 1 \leq j \leq n}} \frac{p_{k_n}}{\prod_{1 \leq l \leq n-1} p(\lambda_{\text{out}}, v_l)} \\ &= \sum_{\substack{k_n: \lambda_{(\text{out},h^*)} \upharpoonright_{\{v_j\}} = \lambda_{(\text{out},k_j)} \upharpoonright_{\{v_j\}} \\ \forall 1 \leq j \leq n}} p_{k_n}. \end{aligned}$$

Notice that

$$\sum_{\substack{k_n: \lambda_{(\text{out}, h^*)} \upharpoonright_{\{v_j\}} = \lambda_{(\text{out}, k_j)} \upharpoonright_{\{v_j\}} \\ \forall 1 \leq j \leq n}} p_{k_n} = p_{h^*}$$

as H_n is necessarily isomorphic to G . By the hypothesis,

$$\sum_{h: \lambda_{(\text{out}, h)} \text{ is valid for } (G, \lambda_{\text{in}})} p_h \geq 1 - \varepsilon,$$

implying that \mathcal{A} succeeds with probability at least $1 - \varepsilon$.

5 Bounded-dependence model can break symmetry

For any graph $G = (V, E)$, a *random process* (or *distribution*) *on the vertices* of G is a family of random variables $\{X_v\}_{v \in V}$, indexed by V , while a *random process on the edges* of G is a family of random variables $\{X_e\}_{e \in E}$ indexed by E . More generally, a random process on G is a family of random variables $\{X_y\}_{y \in V \cup E}$ indexed by $V \cup E$. The variables of a random process live in the same probability space and take values in some label set Σ . In general, we will consider random processes over vertices of graphs unless otherwise specified.

We now introduce the notion of T -dependent distribution. To do so, we extend the definition of distance to edges as follows: For any two edges $e = \{v_1, v_2\}, e' = \{u_1, u_2\} \in E$, $\text{dist}_G(e, e') = \min_{i, j \in [2]} \text{dist}_G(v_i, u_j)$. Similarly, the distance between any edge $e = (v_1, v_2)$ and a vertex v is $\text{dist}_G(e, v) = \min_{i \in [2]} \text{dist}_G(v_i, v)$. The definition extends easily to subsets containing vertices and edges.

Definition 5.1 (T -dependent distribution). Let $T \in \mathbb{N}$ be a natural number and $G = (V, E)$ be any graph. A random process $\{X_v\}_{v \in V}$ on the vertices G is said to be a T -dependent distribution if, for all subsets $S, S' \subseteq V$ such that $\text{dist}_G(S, S') > T$, the two processes $\{X_v\}_{v \in S}$ and $\{X_v\}_{v \in S'}$ are independent. Analogous definitions hold for random processes on the edges of a graph and for random processes on the whole graph.

A way to define T -dependent distributions that uses the same notation of Section 4.1 is by describing the output probability of global labelings: Let I be a discrete set of indices, and $G = (V, E)$ some graph. A distribution $\{(\lambda_i, p_i)\}_{i \in I}$ over output labelings, where $\lambda_i : V \rightarrow \Sigma$ is an output labeling, is T -dependent if the following holds: for all output labelings λ in $\{(\lambda_i, p_i)\}$, for every two subsets of nodes $S_1, S_2 \subseteq V(G)$ such that $\text{dist}_G(S_1, S_2) > T$, we have that

$$p(\lambda, S_1 \cup S_2) = p(\lambda, S_1) \cdot p(\lambda, S_2).$$

Notice that outcomes, as defined in Definition 4.1, output a random process for every input. Often, we have a family of graphs of arbitrarily large size n on which a $T(n)$ -dependent distribution is defined.¹

Now we define the hypothetical computational model that outputs T -dependent distributions on some input graph.

The bounded-dependence model. The bounded-dependence model is a computational model that, for a given family of graphs \mathcal{F} , produces an outcome (as defined in Definition 4.1) over \mathcal{F} that is non-signaling beyond distance $T = T(G, x)$ (as defined in Definition 4.2), where $G \in \mathcal{F}$ and x represents the input, which, in turn, produces $T(G, x)$ -dependent distributions. The random processes produced by an outcome are said to be *finitely-dependent* if $T = O(1)$ for all graphs in \mathcal{F} and all input data. If an outcome with the aforementioned

¹Note that the dependency T might depend on other graph parameters such as the maximum degree Δ , the chromatic number χ , etc.

properties solves a problem Π over a graph family \mathcal{F} with probability at least p , we say that the pair (Π, \mathcal{F}) has complexity T (with success probability p), and T is also called the *locality* or the *complexity* of the corresponding output distributions.

We are particularly interested in T -dependent distributions that satisfy invariance properties: We say that a random process $\{X_v\}_{v \in V}$ over vertices of a graph $G = (V, E)$ is *invariant under automorphisms* if, for all automorphisms $f : V \rightarrow V$ of $G = (V, E)$, the two processes $\{X_v\}_{v \in V}$ and $\{Y_v = X_{f(v)}\}_{v \in V}$ are equal in law. The definition is easily extendable to random processes on edges and random processes on the whole graph.

A stronger requirement is the *invariance under subgraph isomorphism*: Suppose we have an outcome $\mathcal{O} : (G, x) \mapsto \{X_v\}_{v \in V(G)}$ that maps each input graph G from family of graphs \mathcal{F} and any input data x to a $T = T(G, x)$ -dependent distribution over G from a family of random process \mathcal{R} . We say that the random process over vertices in \mathcal{R} are invariant under subgraph isomorphisms if, given any two graphs $G_1, G_2 \in \mathcal{F}$ of size n_1, n_2 with associated process $\{X_v^{(1)}\}_{v \in V(G_1)}$ and $\{X_v^{(2)}\}_{v \in V(G_2)}$, and any two subgraphs $H_1 \subseteq G_1, H_2 \subseteq G_2$ such that $G_1[\mathcal{N}_{T(n_1)}(H_1)]$ and $G_2[\mathcal{N}_{T(n_2)}(H_2)]$ are isomorphic (with the isomorphism that brings H_1 into H_2), then $\{X_v^{(1)}\}_{v \in V(H_1)}$ and $\{X_v^{(2)}\}_{v \in V(H_2)}$ are equal in law.² Trivially, invariance under subgraph isomorphisms implies invariance under automorphisms and the non-signaling property. This definition is again easily extendable to the case of families of random processes over edges or over graphs in general.

The baseline of our result is a finitely-dependent distribution provided by [41, 43] on paths and cycles.

Theorem 5.2 (Finitely-dependent coloring of the integers and of cycles [41, 43]). *Let $G = (V, E)$ be a graph that is either a cycle with at least 2 nodes or has $V = \mathbb{Z}$ and $E = \{\{i, i + 1\} : i \in \mathbb{Z}\}$. For $(k, q) \in \{(1, 4), (2, 3)\}$, there exists a k -dependent distribution $\{X_v^{(G)}\}_{v \in V(G)}$ that gives a q -coloring of G . Furthermore, such distributions can be chosen to meet the following properties: If $\{X_i\}_{i \in [n]}$ is the k -dependent q -coloring of the n -cycle and $\{Y_i\}_{i \in \mathbb{Z}}$ is the k -dependent q -coloring of the integers, then $\{X_i\}_{i \in [n-k]}$ is equal in law to $\{Y_i\}_{i \in [n-k]}$. If $n = 2$, then the finitely-dependent colorings on the 2-cycle and on a path of 2 nodes are identical in distribution.*

It is immediate that the disjoint union of any number of paths and cycles (even countably many) admits such distributions as well.

Corollary 5.3. *Let \mathcal{F} be the family of graphs formed by the disjoint union (possibly uncountably many) paths with countably many nodes and cycles of any finite length. For all $G \in \mathcal{F}$ and for $(k, q) \in \{(1, 4), (2, 3)\}$, there exists a k -dependent distribution $\{X_v^{(G)}\}_{v \in V(G)}$ that gives a q -coloring of G . Furthermore, such distributions can be chosen to meet all the following properties:*

1. *On each connected component $H \subseteq G$, $\{X_v^{(G)}\}_{v \in V(H)}$ is given by Theorem 5.2.*
2. *$\{\{X_v^{(G)}\}_{v \in V(G)}\}_{G \in \mathcal{F}}$ is invariant under subgraph isomorphisms.*

We will use this result to provide “fake local identifiers” to the nodes of the graph to simulate SLOCAL through a random process with constant dependency. In order to do that, we introduce some results on the composition of T -dependent distributions.

Trivially, every T -round (deterministic or randomized) port-numbering algorithm defines a $2T$ -dependent distribution over the input graph. Furthermore, if the underlying port-numbering model has access to random bits, the distribution can be made invariant under subgraph isomorphisms (provided that, whenever a distribution over input labelings is given together with the input graph, such distribution is also invariant under subgraph isomorphisms). The composition of the T_2 -dependent distribution obtained by a T_2 -round port-numbering algorithm and any T_1 -dependent distribution yields a $(2T_2 + T_1)$ -dependent distribution.

Lemma 5.4. *Let $\Sigma^{(1)}$ and $\Sigma^{(2)}$ be two label sets with countably many labels. Let \mathcal{F} be any family of graphs, and let $\mathcal{R} = \{\{X_v\}_{v \in V(G)} : G \in \mathcal{F}\}$ be a family of T_1 -dependent distributions taking values in $\Sigma^{(1)}$, where T_1*

²We remark that, as opposed to Definition 4.2, the isomorphism must preserve the input but *not* the node identifiers.

might depend on parameters of G . Consider any T_2 -round port-numbering algorithm \mathcal{A} that takes as an input $G \in \mathcal{F}$ labelled by $\{X_v\}_{v \in V(G)}$: it defines another distribution $\{Y_v\}_{v \in V(G)}$ taking values in some label set $\Sigma^{(2)}$. Then, the following properties hold:

1. $\{Y_v\}_{v \in V(G)}$ is a $(2T_2 + T_1)$ -dependent distribution on G .
2. If the processes in \mathcal{R} are invariant under subgraph isomorphisms, T_2 is constant, and \mathcal{A} does not depend on the size of the input graph and permutes port numbers locally u.a.r. at round 0, then the processes in $\{\{Y_v\}_{v \in V(G)} : G \in \mathcal{F}\}$ are invariant under subgraph isomorphisms.

Proof. We prove the claim 1 first. Fix any $G = (V, E) \in \mathcal{F}$, and the corresponding T_2 -dependent distribution $\{X_v\}_{v \in V} \in \mathcal{R}$. Fix any two subsets $S, S' \subseteq V$ such that $\text{dist}_G(S, S') > 2T_2 + T_1$. Consider any output labeling $\lambda^{(2)} : V \rightarrow \Sigma^{(2)}$. Then,

$$\begin{aligned}
& \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v = \lambda^{(2)}(v)\} \right] \\
&= \sum_{\lambda^{(1)} : V \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in V} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in V} \{X_v = \lambda^{(1)}(v)\} \right] \\
&= \sum_{\lambda^{(1)} : V \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in V} \{X_v = \lambda^{(1)}(v)\} \right] \\
&= \sum_{\lambda^{(1)} : \mathcal{N}_{T_2}(S \cup S') \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right],
\end{aligned} \tag{3}$$

where Eq. (3) holds because the output of $\{Y_v\}_{v \in S \cup S'}$ is independent of $\{X_v\}_{v \notin \mathcal{N}_{T_2}(S \cup S')}$. Since $\text{dist}_G(S, S') > 2T_2$,

$$\begin{aligned}
& \sum_{\lambda^{(1)} : \mathcal{N}_{T_2}(S \cup S') \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&= \sum_{\lambda^{(1)} : \mathcal{N}_{T_2}(S \cup S') \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&= \sum_{\lambda^{(1)} : \mathcal{N}_{T_2}(S \cup S') \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S)} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
&\quad \cdot \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right].
\end{aligned}$$

Now observe that $\text{dist}_G(\mathcal{N}_{T_2}(S), \mathcal{N}_{T_2}(S')) > T_1$. Since $\{X_v\}_{v \in V}$ is a T_1 -dependent distribution, it holds that

$$\sum_{\lambda^{(1)} : \mathcal{N}_{T_2}(S \cup S') \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S)} \{X_v = \lambda^{(1)}(v)\} \right]$$

$$\begin{aligned}
& \cdot \Pr \left[\bigcap_{v \in S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
& \cdot \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S \cup S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
= & \sum_{\lambda^{(1)}: \mathcal{N}_{T_2}(S \cup S') \rightarrow \Sigma^{(1)}} \Pr \left[\bigcap_{v \in S} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S)} \{X_v = \lambda^{(1)}(v)\} \right] \\
& \cdot \Pr \left[\bigcap_{v \in S'} \{Y_v = \lambda^{(2)}(v)\} \middle| \bigcap_{v \in \mathcal{N}_{T_2}(S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
& \cdot \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S)} \{X_v = \lambda^{(1)}(v)\} \right] \Pr \left[\bigcap_{v \in \mathcal{N}_{T_2}(S')} \{X_v = \lambda^{(1)}(v)\} \right] \\
= & \Pr \left[\bigcap_{v \in S} \{Y_v = \lambda^{(2)}(v)\} \right] \cdot \Pr \left[\bigcap_{v \in S'} \{Y_v = \lambda^{(2)}(v)\} \right].
\end{aligned}$$

We now prove claim 2. Given any $G, H \in \mathcal{F}$ of size n_G and n_H , respectively, consider any subgraph isomorphism $f: \mathcal{N}_{2T_2+T_1}(n_G)(G', G) \rightarrow \mathcal{N}_{2T_2+T_1}(n_H)(H', H)$ for any two $G' \subseteq G, H' \subseteq H$ such that f restricted to G' is an isomorphism to H' . Let $\{X_v^{(G)}\}_{v \in V(G)}$ and $\{X_v^{(H)}\}_{v \in V(H)}$, and $\{Y_v^{(G)}\}_{v \in V(G)}$ and $\{Y_v^{(H)}\}_{v \in V(H)}$ be the corresponding distributions of interest before and after the combination with the port-numbering algorithm, respectively. Fix any subset of nodes $U \subseteq V(G')$ and consider any family of labels $\{\lambda_u\}_{u \in U}$ indexed by U . In order to prove property 1, it is sufficient to show that

$$\Pr \left[\bigcap_{u \in U} \{Y_u^{(G)} = \lambda_u\} \right] = \Pr \left[\bigcap_{u \in U} \{Y_{f(u)}^{(H)} = \lambda_u\} \right].$$

Notice that $\Pr \left[\bigcap_{u \in U} \{Y_u = \lambda_u\} \right]$ depends solely on the graph $G[\mathcal{N}_{T_2}(U, G)] = \cup_{u \in U} G[\mathcal{N}_{T_2}(u, G)]$, the distribution of the port numbers in $G[\mathcal{N}_{T_2}(U, G)]$, and the random process $\{X_u\}_{u \in \mathcal{N}_{T_2}(U)}$. By hypothesis, $\{X_u^{(G)}\}_{u \in \mathcal{N}_{T_2}(U, G)}$ is equal in law to $\{X_{f(u)}^{(H)}\}_{u \in \mathcal{N}_{T_2}(U, G)}$. Furthermore, the restriction of f to $G[\mathcal{N}_{T_2}(U, G)]$ defines an isomorphism from $G[\mathcal{N}_{T_2}(U, G)]$ to $H[\mathcal{N}_{T_2}(f(U), H)]$, hence the distribution of the port numbers in $G[\mathcal{N}_{T_2}(U, G)]$ is the same as that in $H[\mathcal{N}_{T_2}(f(U), H)]$ because each node permutes the port numbers locally u.a.r. Thus, $\Pr \left[\bigcap_{u \in U} \{Y_u^{(G)} = \lambda_u\} \right]$ must be the same as $\Pr \left[\bigcap_{u \in U} \{Y_{f(u)}^{(H)} = \lambda_u\} \right]$. \square

T -dependent distributions over different graphs can be combined to obtain a T -dependent distribution over the graph union.

Lemma 5.5. *Let $\{X_v\}_{v \in V_1}$ and $\{Y_v\}_{v \in V_2}$ be a T_1 -dependent distribution over a graph $G_1 = (V_1, E_1)$ taking values in $\Sigma^{(1)}$ and a T_2 -dependent distribution over a graph $G_2 = (V_2, E_2)$ taking values in $\Sigma^{(2)}$, respectively. Assume $\{X_v\}_{v \in V_1}$ and $\{Y_v\}_{v \in V_2}$ to be independent processes. Consider the graph $H = (V = V_1 \cup V_2, E = E_1 \cup E_2)$ and a distribution $\{Z_v\}_{v \in V}$ over H taking values in $\Sigma = (\Sigma^{(1)} \cup \{0\}) \times (\Sigma^{(2)} \cup \{0\})$ defined by*

$$Z_v = \begin{cases} (X_v, 0) & \text{if } v \in V_1 \setminus V_2, \\ (X_v, Y_v) & \text{if } v \in V_1 \cap V_2, \\ (0, Y_v) & \text{if } v \in V_2 \setminus V_1. \end{cases}$$

Then, $\{Z_v\}_{v \in V}$ is $\max(T_1, T_2)$ -dependent.

Proof. For any vector $\mathbf{v} \in \Sigma_1 \times \cdots \times \Sigma_n$, we write $\mathbf{v}[i]$ to denote its i -th entry. Fix any output labeling $\lambda: V \rightarrow \Sigma$ such that $\lambda(v)[2] = 0$ for all $v \in V_1 \setminus V_2$ and $\lambda(v)[1] = 0$ for all $v \in V_2 \setminus V_1$. Observe that for any subset $S \subseteq V$ it holds that

$$\begin{aligned}
& \Pr \left[\bigcap_{v \in S} \{Z_v = \lambda(v)\} \right] \\
= & \Pr \left[\left(\bigcap_{v \in S \cap V_1} \{X_v = \lambda(v)[1]\} \right) \cap \left(\bigcap_{v \in S \cap V_2} \{Y_v = \lambda(v)[2]\} \right) \right] \\
= & \Pr \left[\bigcap_{v \in S \cap V_1} \{X_v = \lambda(v)[1]\} \right] \cdot \Pr \left[\bigcap_{v \in S \cap V_2} \{Y_v = \lambda(v)[2]\} \right], \tag{4}
\end{aligned}$$

where the latter equality follows by independence between $\{X_v\}_{v \in V_1}$ and $\{Y_v\}_{v \in V_2}$. W.l.o.g., suppose $T_1 \geq T_2$. Consider two subsets $S, S' \subseteq V$ such that $\text{dist}_G(S, S') > T_1$. Fix any output labeling $\lambda : V \rightarrow \Sigma$ such that $\lambda(v)[2] = 0$ for all $v \in V_1 \setminus V_2$ and $\lambda(v)[1] = 0$ for all $v \in V_2 \setminus V_1$. Using Eq. (4), we have that

$$\begin{aligned}
& \Pr[\cap_{v \in S \cup S'} \{Z_v = \lambda(v)\}] \\
&= \Pr[\cap_{v \in (S \cup S') \cap V_1} \{X_v = \lambda(v)[1]\}] \cdot \Pr[\cap_{v \in (S \cup S') \cap V_2} \{Y_v = \lambda(v)[2]\}] \\
&= \Pr[\cap_{v \in (S \cap V_1) \cup (S' \cap V_1)} \{X_v = \lambda(v)[1]\}] \cdot \Pr[\cap_{v \in (S \cap V_2) \cup (S' \cap V_2)} \{Y_v = \lambda(v)[2]\}] \\
&= \Pr[\cap_{S \cap V_1} \{X_v = \lambda(v)[1]\}] \cdot \Pr[\cap_{S' \cap V_1} \{X_v = \lambda(v)[1]\}] \tag{5} \\
&\quad \cdot \Pr[\cap_{S \cap V_2} \{Y_v = \lambda(v)[2]\}] \cdot \Pr[\cap_{S' \cap V_2} \{Y_v = \lambda(v)[2]\}] \\
&= \Pr[(\cap_{v \in S \cap V_1} \{X_v = \lambda(v)[1]\}) \cap (\cap_{v \in S \cap V_2} \{Y_v = \lambda(v)[2]\})] \tag{6} \\
&\quad \cdot \Pr[(\cap_{v \in S' \cap V_1} \{X_v = \lambda(v)[1]\}) \cap (\cap_{v \in S' \cap V_2} \{Y_v = \lambda(v)[2]\})] \\
&= \Pr[\cap_{v \in S} \{Z_v = \lambda(v)\}] \cdot \Pr[\cap_{v \in S'} \{Z_v = \lambda(v)\}],
\end{aligned}$$

where Eq. (5) holds because $\{X_v\}_{v \in V_1}$ is T_1 -dependent and $\{Y_v\}_{v \in V_2}$ is T_2 -dependent, while Eq. (6) holds because $\{X_v\}_{v \in V_1}$ and $\{Y_v\}_{v \in V_2}$ are independent. \square

Now we present a final lemma on the composition of random processes. To do so, we first introduce the notation of *random decomposition* of a graph.

Definition 5.6 (Random decomposition). Let $G = (V, E)$ be any graph and \mathcal{P} a family of subgraphs of G . For any $k \in \mathbb{N}$, let $\Gamma(G)$ be a random variable taking values in \mathcal{P}^k that is sampled according to any probability distribution. We say that $\Gamma(G)$ is a *random k -decomposition* of G in \mathcal{P} .

Given a random k -decomposition $\Gamma(G)$ of G in \mathcal{P} , for any $y \in V \cup E$, we define the random variable $\Gamma(G)_y \in \{0, 1\}^k$ as follows: $\Gamma(G)_y[i] = 1$ if y belongs to $\Gamma(G)[i]$ and 0 otherwise. Notice that $\{\Gamma(G)_y\}_{y \in V \cup E}$ is a random process on G . If $\{\Gamma(G)_y\}_{y \in V \cup E}$ is invariant under automorphisms, then we say that the random k -decomposition $\Gamma(G)$ is invariant under automorphisms. If, for a family of graphs \mathcal{F} and any graph $G \in \mathcal{F}$, $\{\Gamma(G)_y\}_{y \in V(G) \cup E(G)}$ is T -dependent (with T being a function of the size of G) and the processes in $\{\{\Gamma(G)_y\}_{y \in V(G) \cup E(G)} : G \in \mathcal{F}\}$ are invariant under subgraph isomorphisms, then we say that the random k -decompositions in $\{\Gamma(G) : G \in \mathcal{F}\}$ are T -dependent and invariant under subgraph isomorphisms.

For a random decomposition, we define the notion of induced random process.

Definition 5.7 (Induced process). Let $G = (V, E)$ be a graph that admits a family of subgraphs \mathcal{P} . Suppose there exists a random process $\{X_v\}_{v \in V(H(\mathcal{P}))}$ with $H(\mathcal{P})$ being the graph obtained by the disjoint union of all elements of \mathcal{P} . Let $\Gamma(G)$ be a random k -decomposition of G in \mathcal{P} . For all $v \in V$ and $G' \in \mathcal{P}$, define the random process $\{X_v^{(G')}\}_{v \in V}$ by setting $X_v^{(G')} = X_{f_{G'}(v)}$ for all $v \in V(G')$, where $f_{G'} : V(G') \rightarrow V(H(\mathcal{P}))$ is the natural immersion of G' into $H(\mathcal{P})$ otherwise set $X_v^{(G')} = 0$. Let $\{Y_v^{(\Gamma(G))}\}_{v \in V}$ be a random process that we define conditional on the output of $\Gamma(G)$: for all $\mathbf{G} \in \mathcal{P}^k$, conditional on $\Gamma(G) = \mathbf{G}$,

$$Y_v^{(\Gamma(G))} = Y_v^{(\mathbf{G})} = (X_v^{(\mathbf{G}[1])}, \dots, X_v^{(\mathbf{G}[k])}).$$

The random process $\{Y_v^{(\Gamma(G))}\}_{v \in V}$ on G is said to be induced by the action of $\{X_v\}_{v \in V(H(\mathcal{P}))}$ over the random k -decomposition $\Gamma(G)$.

Now we present a result on the induced random process whenever the random decomposition and the family of random processes that acts on the random decomposition meet some invariance properties.

Lemma 5.8. *Let \mathcal{F} be a family of graphs. For any $G = (V, E) \in \mathcal{F}$, let \mathcal{P}_G be any family of subgraphs of G that is closed under node removal and disjoint graph union, that is, if $G_1, G_2 \in \mathcal{P}_G$ and $V(G_1) \cap V(G_2) =$*

\emptyset , then $G_1 \cup G_2 \in \mathcal{P}_G$. Furthermore, suppose that, for each pair of isomorphic subgraphs $G_1, G_2 \subseteq G$, $G_1 \in \mathcal{P}_G \implies G_2 \in \mathcal{P}_G$. Moreover, let $\Gamma(G)$ be a random k -decomposition of G in \mathcal{P}_G that is T_1 -dependent, and suppose that the random decompositions in $\{\Gamma(G) : G \in \mathcal{F}\}$ are invariant under subgraph isomorphism. Suppose there is T_2 -dependent distribution $\{X_v\}_{v \in V(H(\mathcal{P}_G))}$, taking values in a finite set Σ , such that $\{\{X_v\}_{v \in V(H(\mathcal{P}_G))} : G \in \mathcal{F}\}$ is invariant under subgraph isomorphism, where $H(\mathcal{P}_G)$ is obtained by the disjoint union of a copy of each element of \mathcal{P}_G . Let $\{Y_v^{\Gamma(G)}\}_{v \in V(G)}$ be the random process induced by the action of $\{X_v\}_{v \in V(H(\mathcal{P}_G))}$ over $\Gamma(G)$. Then, $\{Y_v^{\Gamma(G)}\}_{v \in V(G)}$ is a $(T_1 + 2T_2)$ -dependent distribution. Furthermore, the processes in $\{\{X_v\}_{v \in V(H(\mathcal{P}_G))} : G \in \mathcal{F}\}$ are invariant under subgraph isomorphism.

Proof. Fix $G \in \mathcal{F}$ and let $\Gamma = \Gamma(G)$, $\mathcal{P} = \mathcal{P}_G$. Since \mathcal{P}^k might be uncountable, we consider the density function f_Γ and the probability measure \mathbb{P} associated to Γ . Consider two subsets of nodes $S, S' \subseteq V$ at distance at least $\max(T_1, T_2) + 1$. Fix any labeling $\lambda : V \rightarrow \Sigma^k$. It holds that

$$\begin{aligned} & \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v^{\Gamma} = \lambda(v)\} \right] \\ &= \int_{\mathcal{P}^k} \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v^{\Gamma} = \lambda(v)\} \mid \Gamma = \mathbf{G} \right] f_\Gamma(\mathbf{G}) \, d\mathbb{P} \end{aligned} \quad (7)$$

$$\begin{aligned} &= \int_{\mathcal{P}^k} \Pr \left[\bigcap_{v \in S \cup S'} \{Y_v^{\mathbf{G}} = \lambda(v)\} \right] f_\Gamma(\mathbf{G}) \, d\mathbb{P} \\ &= \int_{\mathcal{P}^k} \Pr \left[\bigcap_{v \in S} \{Y_v^{\mathbf{G}} = \lambda(v)\} \right] \Pr \left[\bigcap_{v \in S'} \{Y_v^{\mathbf{G}} = \lambda(v)\} \right] f_\Gamma(\mathbf{G}) \, d\mathbb{P} \end{aligned} \quad (8)$$

$$= \int_{\mathcal{P}^k} \prod_{i \in [k]} \Pr \left[\bigcap_{v \in S} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\} \right] \Pr \left[\bigcap_{v \in S'} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\} \right] f_\Gamma(\mathbf{G}) \, d\mathbb{P}, \quad (9)$$

where Eq. (7) holds by the law of total probability, Eq. (8) holds since $\{Y_v^{\mathbf{G}}\}_{v \in V}$ is T_2 -dependent for all $\mathbf{G} \in \mathcal{P}^k$ by Lemma 5.5, and Eq. (9) holds since $\{X_v^{\mathbf{G}[i]}\}_{v \in V}$ and $\{X_v^{\mathbf{G}[j]}\}_{v \in V}$ are independent if $i \neq j$. Notice that, for any set $S \subseteq V$, $G' = \mathbf{G}[i][\mathcal{N}_{T_2}(S, \mathbf{G}[i])] \in \mathcal{P}$ and the event $\bigcap_{v \in S} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\}$ has the same probability as $\bigcap_{v \in S} \{X_v^{G'} = \lambda(v)[i]\}$ because $\{X_v\}_{v \in V(H(\mathcal{P}))}$ is invariant under subgraph isomorphisms.

For any subset $U \subseteq V$ and any integer $T \geq 0$, let

$$\mathcal{H}(T, U) = \left\{ (\mathbf{G}[1][\mathcal{N}_T(U, \mathbf{G}[1])], \dots, \mathbf{G}[k][\mathcal{N}_T(U, \mathbf{G}[k])]) \mid \mathbf{G} \in \mathcal{P}^k \right\} \subseteq \mathcal{P}^k,$$

and let $\mathbb{P}[\mathcal{H}(T, U)]$ be the restriction of \mathbb{P} to $\mathcal{H}(T, U)$ defined as follows: for any event E ,

$$\mathbb{P}[\mathcal{H}(T, U)](E) = \mathbb{P}(E \cap \mathcal{H}(T, U)) / \mathbb{P}(\mathcal{H}(T, U)).$$

Finally, define the random variable $\Gamma^{(T, U)}$ to be a k -dimensional vector, taking values in $\mathcal{H}(T, U)$, whose i -th entry is $\Gamma[i][\mathcal{N}_T(U, \Gamma[i])]$, and let $f_{\Gamma^{(T, U)}}$ be its density function. Define $U = S \cup S'$: Eq. (9) becomes

$$\begin{aligned} & \int_{\mathcal{P}^k} \prod_{i \in [k]} \Pr \left[\bigcap_{v \in S} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\} \right] \Pr \left[\bigcap_{v \in S'} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\} \right] f_\Gamma(\mathbf{G}) \, d\mathbb{P} \\ &= \int_{\mathcal{H}(T_2, U)} \prod_{i \in [k]} \Pr \left[\bigcap_{v \in S} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\} \right] \Pr \left[\bigcap_{v \in S'} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\} \right] f_{\Gamma^{(T_2, U)}}(\mathbf{H}) \, d\mathbb{P}[\mathcal{H}(T_2, U)]. \end{aligned} \quad (10)$$

Now notice that, since the random decomposition Γ is T_1 -dependent and $\text{dist}_G(S, S') > T_1 + 2T_2$, the probability space $\mathcal{H}(T_2, U)$ (with measure $\mathbb{P}[\mathcal{H}(T_2, U)]$) is isomorphic to the product space $\mathcal{H}(T_2, S) \times \mathcal{H}(T_2, S')$ (with product measure $\mathbb{P}[\mathcal{H}(T_2, S)] \times \mathbb{P}[\mathcal{H}(T_2, S')]$). The isomorphism brings any element

$$(\mathbf{G}[1][\mathcal{N}_{T_2}(U, \mathbf{G}[1])], \dots, \mathbf{G}[k][\mathcal{N}_{T_2}(U, \mathbf{G}[k])])$$

of $\mathcal{H}(T_2, U)$ into

$$((\mathbf{G}[1][\mathcal{N}_{T_2}(S, \mathbf{G}[1])], \dots, \mathbf{G}[k][\mathcal{N}_{T_2}(S, \mathbf{G}[k])]), (\mathbf{G}[1][\mathcal{N}_{T_2}(S', \mathbf{G}[1])], \dots, \mathbf{G}[k][\mathcal{N}_{T_2}(S', \mathbf{G}[k])]))$$

which belongs to $\mathcal{H}(T_2, S) \times \mathcal{H}(T_2, S')$. Hence, we get

$$\begin{aligned} & \int_{\mathcal{H}(T_2, U)} \prod_{i \in [k]} \Pr[\cap_{v \in S} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] \Pr[\cap_{v \in S'} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] f_{\Gamma(T_2, U)}(\mathbf{H}) d\mathbb{P}[\mathcal{H}(T_2, S)] \\ &= \int_{\mathcal{H}(T_2, S)} \int_{\mathcal{H}(T_2, S')} \prod_{i \in [k]} \Pr[\cap_{v \in S} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] \Pr[\cap_{v \in S'} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] \\ & \quad \cdot f_{\Gamma(T_2, S)}(\mathbf{H}_1) f_{\Gamma(T_2, S')}(\mathbf{H}_2) d\mathbb{P}[\mathcal{H}(T_2, S)] d\mathbb{P}[\mathcal{H}(T_2, S')]. \end{aligned}$$

The latter becomes

$$\begin{aligned} & \int_{\mathcal{H}(T_2, S)} \int_{\mathcal{H}(T_2, S')} \prod_{i \in [k]} \Pr[\cap_{v \in S} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] \Pr[\cap_{v \in S'} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] \\ & \quad \cdot f_{\Gamma(T_2, S)}(\mathbf{H}_1) f_{\Gamma(T_2, S')}(\mathbf{H}_2) d\mathbb{P}[\mathcal{H}(T_2, S)] d\mathbb{P}[\mathcal{H}(T_2, S')] \\ &= \int_{\mathcal{H}(T_2, S)} \prod_{i \in [k]} \Pr[\cap_{v \in S} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] f_{\Gamma(T_2, S)}(\mathbf{H}_1) d\mathbb{P}[\mathcal{H}(T_2, S)] \\ & \quad \cdot \int_{\mathcal{H}(T_2, S')} \prod_{i \in [k]} \Pr[\cap_{v \in S'} \{X_v^{\mathbf{H}[i]} = \lambda(v)[i]\}] f_{\Gamma(T_2, S')}(\mathbf{H}_2) d\mathbb{P}[\mathcal{H}(T_2, S')] \\ &= \int_{\mathcal{P}^k} \prod_{i \in [k]} \Pr[\cap_{v \in S} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\}] f_{\Gamma}(\mathbf{G}) d\mathbb{P} \\ & \quad \cdot \int_{\mathcal{P}^k} \prod_{i \in [k]} \Pr[\cap_{v \in S'} \{X_v^{\mathbf{G}[i]} = \lambda(v)[i]\}] f_{\Gamma}(\mathbf{G}) d\mathbb{P} \\ &= \Pr[\cap_{v \in S} \{Y_v^{(\Gamma(G))} = \lambda(v)\}] \Pr[\cap_{v \in S'} \{Y_v^{(\Gamma(G))} = \lambda(v)\}], \end{aligned} \tag{11}$$

where Eq. (11) follows by the same reasoning as for Eq. (10) but in reverse.

Now we want to prove that $\{\{Y_v^{(\Gamma(G))}\}_{v \in V} : G \in \mathcal{F}\}$ is invariant under subgraph isomorphism. Fix any two graphs $G, H \in \mathcal{F}$ of sizes n_G and n_H , respectively. Consider any isomorphism α between the radius- $(T_1(n_G) + 2T_2(n_G))$ neighborhoods of any subgraph $G' \subseteq G$ and the radius- $(T_1(n_H) + 2T_2(n_H))$ neighborhoods of any subgraph $H' \subseteq H$, such that the restriction of α to G' is an isomorphism to H' . With an abuse of notation, for any subgraph $K \subseteq G'$, let us denote $\alpha(K) \subseteq H'$ its isomorphic image in H' through α . Let $T_G = T_1(n_G) + 2T_2(n_G)$ and $T_H = T_1(n_H) + 2T_2(n_H)$. Fix any labeling $\lambda : V \rightarrow \Sigma^k$. We have that

$$\begin{aligned} & \Pr[\cap_{v \in V(G')} \{Y_v^{(\Gamma(G))} = \lambda(v)\}] \\ &= \int_{\mathcal{P}_G^k} \Pr[\cap_{v \in V(G')} \{Y_v^{(\Gamma(G))} = \lambda(v)\} \mid \Gamma(G) = \mathbf{G}] f_{\Gamma(G)}(\mathbf{G}) d\mathbb{P}_G. \\ &= \int_{\mathcal{P}_G^k} \Pr[\cap_{v \in V(G')} \{Y_v^{(\mathbf{G})} = \lambda(v)\}] f_{\Gamma(G)}(\mathbf{G}) d\mathbb{P}_G \\ &= \int_{\mathcal{P}_G^k} \prod_{i \in [k]} \Pr[\cap_{v \in V(G')} \{X_v^{(\mathbf{G}[i])} = \lambda(v)[i]\}] f_{\Gamma(G)}(\mathbf{G}) d\mathbb{P}_G \\ &= \int_{\mathcal{H}_G(T_G, V(G'))} \prod_{i \in [k]} \Pr[\cap_{v \in V(G')} \{X_v^{(\mathbf{H}[i])} = \lambda(v)[i]\}] f_{\Gamma(G)(T_G, V(G'))}(\mathbf{H}) d\mathbb{P}_G(\mathcal{H}_G(T_G, V(G'))) \end{aligned}$$

$$= \int_{\mathcal{H}_G(T_G, V(G'))} \prod_{i \in [k]} \Pr \left[\bigcap_{v \in V(G')} \{X_{\alpha(v)}^{(\alpha(\mathbf{H}[i]))} = \lambda(v)[i]\} \right] f_{\Gamma(G)(T_G, V(G'))}(\mathbf{H}) \, d\mathbb{P}_G(\mathcal{H}_G(T_G, V(G'))),$$

where the latter holds because

$$\Pr \left[\bigcap_{v \in V(G')} \{X_v^{(\mathbf{H}[i])} = \lambda(v)[i]\} \right] = \Pr \left[\bigcap_{v \in V(G')} \{X_{\alpha(v)}^{(\alpha(\mathbf{H}[i]))} = \lambda(v)[i]\} \right]$$

as $\{\{X_v\}_{v \in V(H(\mathcal{P}_G))} : G \in \mathcal{F}\}$ is T_2 -dependent and invariant under subgraph isomorphism. Furthermore, since the processes in $\{\Gamma(G) : G \in \mathcal{F}\}$ are T_1 -dependent and invariant under subgraph isomorphism, we have that $f_{\Gamma(G)(T_G, V(G'))}(\mathbf{H}) = f_{\Gamma(H)(T_H, V(H'))}(\alpha(\mathbf{H}[1]), \dots, \alpha(\mathbf{H}[k]))$ almost everywhere in $\mathcal{H}_G(T_G, V(G'))$, and that the probability space $\mathcal{H}_G(T_G, V(G'))$ with measure $d\mathbb{P}_G(\mathcal{H}_G(T_G, V(G')))$ is isomorphic to $\mathcal{H}_H(T_H, V(H'))$ with measure $d\mathbb{P}_H(\mathcal{H}_H(T_H, V(H')))$, where the isomorphism brings \mathbf{H} into $(\alpha(\mathbf{H}[1]), \dots, \alpha(\mathbf{H}[k]))$. Hence,

$$\begin{aligned} & \int_{\mathcal{H}_G(T_G, V(G'))} \prod_{i \in [k]} \Pr \left[\bigcap_{v \in V(G')} \{X_{\alpha(v)}^{(\alpha(\mathbf{H}[i]))} = \lambda(v)[i]\} \right] f_{\Gamma(G)(T_G, V(G'))}(\mathbf{H}) \, d\mathbb{P}_G(\mathcal{H}_G(T_G, V(G'))) \\ &= \int_{\mathcal{H}_H(T_H, V(H'))} \prod_{i \in [k]} \Pr \left[\bigcap_{v \in V(H')} \{X_v^{(\mathbf{H}[i])} = \lambda(v)[i]\} \right] f_{\Gamma(H)(T_H, V(H'))}(\mathbf{H}) \, d\mathbb{P}_H(\mathcal{H}_H(T_H, V(H'))) \\ &= \Pr \left[\bigcap_{v \in V(H')} \{Y_v^{(\Gamma(H))} = \lambda(v)\} \right], \end{aligned}$$

concluding the proof. □

Remark 5.9. When the underlying graph is a directed graph, Lemma 5.8 guarantees invariance under subgraph isomorphisms that keep edge orientation.

We are going to work on rooted pseudotrees and pseudoforests. A *pseudotree* is a graph that is connected and contains at most one cycle. A *pseudoforest* is a graph obtained by the disjoint union of pseudotrees; an equivalent definition of pseudoforest is a graph in which each connected component has no more edges than vertices. Note that a pseudotree might contain multiple edges: however, we assume it does not contain self-loops as self-loops are useless communication links in the LOCAL model. A *rooted tree* is a tree where each edge is oriented and all nodes have outdegree at most 1: it follows that all but one node have outdegree exactly 1 and one node (the *root*) has outdegree 0. Trivially, a tree can be rooted by selecting one node and orienting all edges towards it. A *rooted pseudotree* is a pseudotree where each edge is oriented and each node has outdegree at most 1: if the pseudotree contains a cycle, then all nodes necessarily have outdegree exactly 1. Any pseudotree can be oriented so that it becomes rooted: just orient the cycle first (if it exists) in a consistent way, then remove it, and make the remaining trees rooted at nodes that belonged to the cycle. A *rooted pseudoforest* is the union of rooted pseudotrees.

We will show that pseudoforests of maximum degree Δ admit a $O(\log^* \Delta)$ -dependent 3-coloring distributions. In order to do so, we use a color reduction technique that follows by known revisions of the Cole–Vishkin technique [28, 39].

Lemma 5.10 (port-numbering algorithm for color reduction in pseudoforests [28, 39]). *Let G be a pseudoforest with countably many nodes. Assume G is given as an input a k -coloring for some $k \geq 3$. There exists a deterministic port-numbering algorithm that does not depend on the size of G and outputs a 3-coloring of G in time $O(\log^* k)$.*

Now we are ready to prove our result on pseudoforests.

Lemma 5.11 (Finitely-dependent coloring of rooted pseudoforests). *Let \mathcal{F} be a family of rooted pseudoforests of countably many nodes of maximum degree Δ . Then, there exists an outcome that associates to each graph $G \in \mathcal{F}$ a $O(\log^* \Delta)$ -dependent distribution on the vertices of G that gives a 3-coloring of G . Furthermore, the family of distributions outputted by the outcome are invariant under subgraph isomorphisms.*

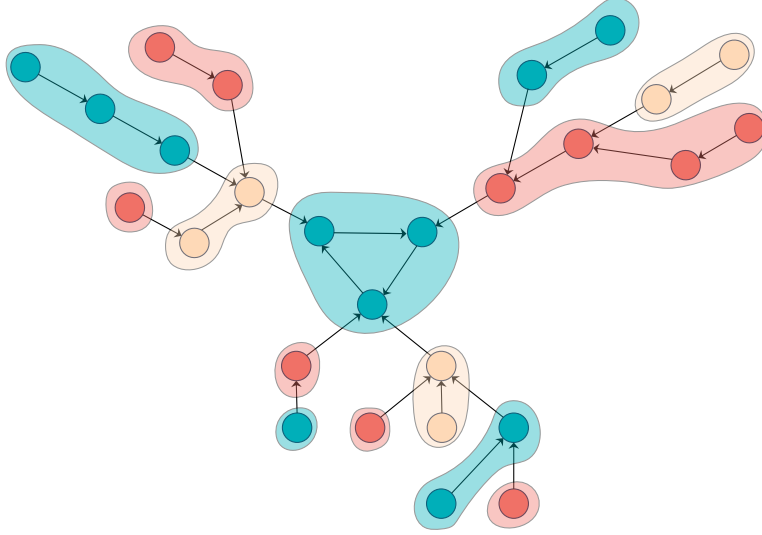


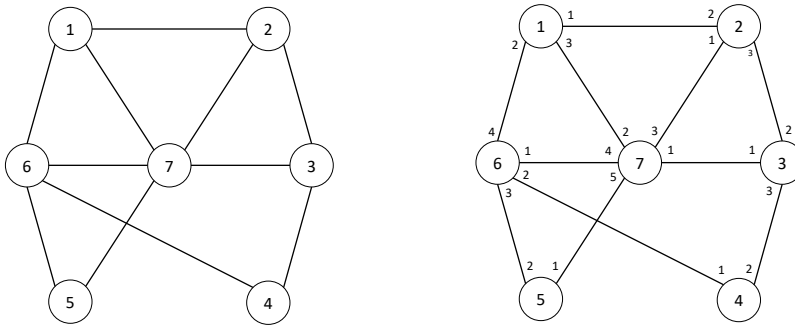
Figure 2: A decomposition of rooted pseudoforests in directed paths and cycles: each node v colors its indegree neighbors with a uniformly sampled permutation of the elements in $[\text{indeg}(v)]$. The graph induced by nodes colored with color i is a disjoint union of directed paths and cycles.

Proof. Let us fix the rooted pseudoforest $G \in \mathcal{F}$. Let \mathcal{P}_G be the family of all subgraphs of G formed by the disjoint union of directed paths and cycles. Notice that \mathcal{P}_G is closed under node removal and disjoint graph union. Furthermore, for any two pair of isomorphic subgraphs $G_1, G_2 \subseteq G$, $G_1 \in \mathcal{P}_G \implies G_2 \in \mathcal{P}_G$. Let $H(\mathcal{P}_G)$ be the graph formed by the disjoint union of a copy of each element of \mathcal{P}_G . By Corollary 5.3, $H(\mathcal{P}_G)$ admits a 1-dependent 4-coloring distribution $\{X_v\}_{v \in H(\mathcal{P}_G)}$ (with colors in $[4]$), such that $\{\{X_v\}_{v \in H(\mathcal{P}_G)} : G \in \mathcal{F}\}$ is invariant under subgraph isomorphism.

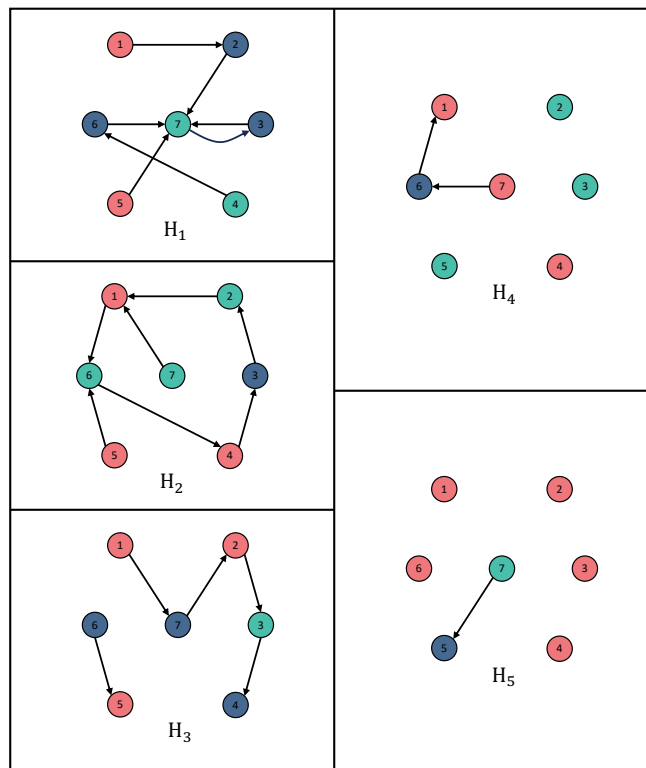
Consider now a (non-proper) coloring of the pseudoforest in which each node u colors its indegree neighbors with a permutation of $\{1, \dots, \text{indeg}(u)\}$ sampled uniformly at random: if a node has outdegree zero, then it is deterministically colored with color 1. Such a coloring is described by a 2-dependent distribution $\{Z_v\}_{v \in V}$ that is (trivially) invariant under subgraph isomorphisms that keep edge orientation. Also, $\{Z_v\}_{v \in V}$ identifies Δ_{in} disjoint random subset of nodes $V_1, \dots, V_{\Delta_{\text{in}}}$, where nodes in V_i are colored with the color i , and Δ_{in} is the maximum indegree of the graph. Let $\Gamma(G) = (G[V_1], \dots, G[V_{\Delta_{\text{in}}}]$, where $\Gamma(G)[i]$ is the random graph induced by V_i . Furthermore, observe that, the output of $\Gamma(G)[i]$ is the disjoint union of oriented paths and/or oriented cycles, with $\Gamma(G)[i]$ being the i -th entry of the Δ_{in} -tuple $\Gamma(G)$ (see Fig. 2). Notice that process $\Gamma(G)$ is 2-dependent and is a random Δ_{in} -decomposition of G in \mathcal{P}_G (according to Definition 5.6), such that the random decompositions in $\{\Gamma(G) : G \in \mathcal{F}\}$ are invariant under subgraph isomorphism. By Lemma 5.8, the random process $\{Y_v^{\Gamma(G)}\}_{v \in V}$, that is induced by the action of $\{X_v\}_{v \in H(\mathcal{P}_G)}$ over the random Δ_{in} -decomposition $\Gamma(G)$ (according to Definition 5.7), is a 4-dependent distribution that gives a $4\Delta_{\text{in}}$ -coloring of G : in fact, $\Gamma(G)[i]$ and $\Gamma(G)[j]$ are disjoint if $i \neq j$, hence only one entry of $Y_v^{\Gamma(G)}$ is non-zero, for all $v \in V$.

We then combine $\{Y_v^{\Gamma(G)}\}_{v \in V}$ and a modified version of the port-numbering algorithm from Lemma 5.10 where at round 0 each node permutes port numbers locally u.a.r.: by Lemma 5.4, we obtain an $O(\log^* \Delta)$ -dependent 3-coloring distribution $\{Q_v^{(G)}\}_{v \in V(G)}$ of G such that processes in $\{\{Q_v^{(G)}\}_{v \in V(G)} : G \in \mathcal{F}\}$ are invariant under subgraph isomorphisms. \square

Our finitely-dependent coloring of pseudoforests can be used as a baseline to provide a $(\Delta + 1)$ -coloring of graphs with maximum degree Δ . The tool we use is again an application of the Cole–Vishkin color reduction technique [55].



(a) Each node v rearranges its port numbers uniformly at random.



(b) The $\Delta = 5$ rooted pseudoforests that we take by considering port numbers $i \in [\Delta]$.

Figure 3: A decomposition of a graph of maximum degree $\Delta = 5$ in rooted pseudoforests: for the sake of image clarity, we focus on the undirected case. In Fig. 3a, each node v rearranges its port-numbers with a uniformly sampled permutation of the elements in $[\text{deg}(v)]$. As shown in Fig. 3b, edges hosting port number i at some endpoint are oriented away from that port (in case both endpoints host port number i , the edge is duplicated) and form a rooted pseudoforest.

Lemma 5.12 (port-numbering algorithm for color reduction of general graphs [55]). *Let $G = (V, E)$ be a graph with maximum degree Δ and countably many nodes. Suppose G is given in input a k -coloring for some $k \geq \Delta + 1$. There exists a deterministic port-numbering algorithm that does not depend on the size of the input graph and outputs a $(\Delta + 1)$ -coloring of G in time $O(\log^* k + \Delta^2)$.*

Lemma 5.13 (Finitely-dependent coloring of bounded-degree graphs). *Let \mathcal{F} be a family of graphs of countably many nodes and maximum degree Δ . Then, there exists an outcome that associates to each graph $G \in \mathcal{F}$ an $O(\Delta^2)$ -dependent distribution on the vertices of G that gives a $(\Delta + 1)$ -coloring of G . Furthermore, the family of distributions outputted by the outcome are invariant under subgraph isomorphisms.*

Proof. Let us fix $G = (V, E) \in \mathcal{F}$. First, if G is not directed, then duplicate each edge and give to each pair of duplicates different orientations. Since a coloring of the original graph is a proper coloring if and only if the same coloring is proper in the directed version, w.l.o.g., we can assume G to be directed.

Let \mathcal{P}_G be a family of all subgraphs of G formed by rooted pseudotrees and disjoint union of rooted pseudotrees. Notice that \mathcal{P}_G is closed under node removal and disjoint graph union. Furthermore, for any two pair of isomorphic subgraphs $G_1, G_2 \subseteq G$, $G_1 \in \mathcal{P}_G \implies G_2 \in \mathcal{P}_G$. The graph $H(\mathcal{P}_G)$ that is the disjoint union of all elements of copies of each \mathcal{P}_G is a rooted pseudoforest of maximum degree Δ and, by Lemma 5.11, admits a 3-coloring $O(\log^* \Delta)$ -dependent distribution $\{X_v\}_{v \in H(\mathcal{P}_G)}$ such that processes in $\{\{X_v\}_{v \in H(\mathcal{P}_G)} : G \in \mathcal{F}\}$ are invariant under subgraph isomorphism.

Now, consider a process in which each node v samples uniformly at random a permutation of a port numbering from $\{1, \dots, \text{outdeg}(v)\}$ for its outdegree edges. For each $i \in [\Delta_{\text{out}}]$, consider the graph G_i induced by edges that host port i : notice that G_i is a rooted pseudoforest as each node has at most one outdegree edge with port i . If a node has degree 0, it deterministically joins G_1 , which remains a pseudoforest. The random choice of port numbering defines a random variable $\Gamma \in \mathcal{P}_G^k$, where $\Gamma[i]$ is the graph induced by port number i : according to Definition 5.6, we obtain a random Δ_{out} -decomposition $\Gamma(G)$ of G in \mathcal{P}_G which is 2-dependent (by construction): also, the random decompositions in $\{\Gamma(G) : G \in \mathcal{F}\}$ are invariant under subgraph isomorphisms. For an example of a possible output of the random decomposition, see Fig. 3.

Hence, the random process $\{Y_v^{(\Gamma(G))}\}_{v \in V}$ from Definition 5.7 is well defined, and provides a proper 3^{Δ} -coloring of G . By Lemma 5.8, the random process $\{Y_v^{(\Gamma(G))}\}_{v \in V}$ is $O(\log^*)$ -dependent and, when $G \in \mathcal{F}$ varies, the processes $\{Y_v^{(\Gamma(G))}\}_{v \in V}$ are invariant under subgraph isomorphisms.

We then combine $\{Y_v^{(\Gamma(G))}\}_{v \in V}$ and a modified version of the port-numbering algorithm from Lemma 5.12 where at round 0 each node permutes port numbers locally u.a.r.: by Lemma 5.4, we obtain an $O(\Delta^2)$ -dependent $(\Delta + 1)$ -coloring distribution of G that is invariant under subgraph isomorphisms (as $G \in \mathcal{F}$ varies). \square

Consider any graph $G = (V, E)$. For any $k \in \mathbb{N}_+$, a distance- k coloring of G is an assignment of colors $c : V \rightarrow \Sigma$ such that, for each node $v \in V$, all nodes in $\mathcal{N}_k(v) \setminus \{v\} = \{u \in V : \text{dist}_G(u, v) \leq k\} \setminus \{v\}$ have colors that are different from $c(v)$.

It is well known that any LCL problem Π that has complexity $O(\log^* n)$ in the LOCAL model has the following property: there exists a constant $k \in \mathbb{N}_+$ (that depends only on the hidden constant in $O(\log^* n)$) such that, if the input graph is given a distance- k coloring, then Π is solvable in time $O(1)$ in the port-numbering model [19]. Furthermore, no knowledge of the size of the input graph is required.

Finding a distance- k coloring in SLOCAL trivially requires locality exactly k , implying the well known fact that LCLs of complexity $O(\log^* n)$ in the LOCAL model have complexity $O(1)$ in SLOCAL. In SLOCAL is indeed easier to address the case of infinite graphs for such problems.

Theorem 5.14. *Consider any LCL problem Π with checking radius r that has complexity $T \geq r$ in the SLOCAL model over a family \mathcal{F} of graphs with maximum degree Δ , where $T = O(1)$ is a constant. For each $G \in \mathcal{F}$, there exists an $O(T^2 \Delta^{2T})$ -dependent distribution $\{Y_v^{(G)}\}_{v \in V(G)}$ that solves Π over G . Furthermore, the processes in $\{\{Y_v^{(G)}\}_{v \in V(G)} : G \in \mathcal{F}\}$ are invariant under subgraph isomorphisms.*

Proof. Fix $G \in \mathcal{F}$ of size n . Consider the power graph G^k , where $k \geq T$, defined by $G^k = \{V, E^k\}$ with $E^k = \{\{u, v\} : u, v \in V, \text{dist}_G(u, v) \leq k\}$. The maximum degree of G^k is Δ^k . By Lemma 5.13, G^k admits an $O(\Delta^{2k})$ -dependent $(\Delta^k + 1)$ -coloring distribution $\{X_v\}_{v \in V}$ that is invariant under subgraph isomorphisms (as $G \in \mathcal{F}$ varies). Notice that such a coloring provides a distance- k coloring for G .

Let \mathcal{A} be the algorithm in SLOCAL that solves Π in time T . Consider a port-numbering algorithm \mathcal{A}' that simulates \mathcal{A} and is defined as follows: At round 0, \mathcal{A}' permutes ports locally u.a.r. Then, nodes colored with color i perform T rounds of communication from round $(i - 1)T + 1$ to round iT and output a label as if they were the i -th node shown by the adversary to the SLOCAL algorithm and had identifier i : all other communications are useless for such nodes. Notice that, for each $v \in V(G)$, identifiers in $\mathcal{N}_T(v)$ are unique and belong to $\{1, \dots, \Delta^k\}$. The running time of \mathcal{A}' is kT . Now we prove that the \mathcal{A}' outputs a correct labeling by induction on the colors of the nodes.

Let us focus on nodes colored with color 1 and, by contradiction, suppose one of such nodes (node u) fails: a failure means that $G[\mathcal{N}_r(u)]$ contains a non-valid labeling. Notice that no node other than u outputs anything in $G[\mathcal{N}_r(u)]$, as nodes in $\mathcal{N}_r(u) \setminus \{u\}$ have color greater than 1. Then \mathcal{A} fails in G where the adversary picks u as the first node in G , reaching a contradiction.

Suppose $c > 1$ and that nodes with color $i < c$ output a correct label. Let us focus on nodes with color c and assume, by contradiction, that one of such nodes (node u) fails: a failure means that $G[\mathcal{N}_r(u)]$ contains a non-valid labeling. Notice that in $G[\mathcal{N}_r(u)]$ only nodes with colors strictly smaller than c have already outputted a label, and such a label is correct by the inductive hypothesis. Furthermore, colors in $\mathcal{N}_r(u)$ are all different as $r \leq k$. Then \mathcal{A} fails in G where the adversary picks nodes in $\mathcal{N}_k(u)$ according to the order given by the coloring, and then continues outside $\mathcal{N}_k(u)$ following any arbitrary order, reaching a contradiction.

Notice that $\{Y_v\}_{v \in V}$, the random process induced by combining $\{X_v\}_{v \in V}$ and \mathcal{A} is a $O(kT\Delta^{2k})$ -dependent distribution on G by Lemma 5.4 with the required invariance properties, and we get the thesis by choosing $k = T$. \square

Theorem 5.14 answers an open question formulated by Holroyd [40].

Corollary 5.15. *Let $G = (V, E)$ be the infinite d -regular trees. There exists a finitely-dependent distribution giving a $(d + 1)$ -coloring of G that is invariant under automorphisms.*

Proof. Finding a $(d + 1)$ -coloring of any d -regular trees has complexity $O(1)$ in SLOCAL: the coloring can just be performed greedily. Theorem 5.14 yields the desired result. \square

6 Simulation of online-LOCAL in SLOCAL for rooted trees

In this section, we show how to turn an online-LOCAL algorithm solving an LCL problem in rooted forests into a deterministic SLOCAL algorithm solving the same problem. More concretely, we prove the following theorem:

Theorem 6.1. *Let Π be an LCL problem with degree constraint Δ , input label set Σ_{in} , output label set Σ_{out} , and checking radius $r > 0$. In addition, let \mathcal{A} be an online-LOCAL algorithm solving Π with locality $T(n)$ over rooted forests. Then the following holds:*

1. *If \mathcal{A} is deterministic, then there exists a deterministic SLOCAL algorithm solving Π with locality $O(r) + T(2^{O(n^3)})$.*
2. *If \mathcal{A} is randomized and has success probability $p(n) > 0$, then there exists a deterministic SLOCAL algorithm solving Π with locality $O(r) + T(2^{O(n^3)} + 2^{O(2n^2)} \cdot \log \frac{1}{p(n)})$.*

Theorem 1.5 follows as a simple corollary of Theorem 6.1, along with the fact that an algorithm with locality $o(\log \log n)$ in the SLOCAL model implies the existence of a locality $o(\log n)$ algorithm in the LOCAL model [36, Proposition III.2].

6.1 Amnesiac online-LOCAL algorithms

We start by formalizing what an online-LOCAL algorithm sees when run for a fixed number of steps on a graph. We then define formally what we mean by amnesiac algorithms:

Definition 6.2 (Partial online-LOCAL run of length ℓ). Let G be a graph with an ordering of nodes v_1, v_2, \dots, v_n . Consider the subgraph $G_\ell \subseteq G$ induced by the radius- T neighborhoods of the first ℓ nodes v_1, \dots, v_ℓ . We call $(G_\ell, (v_1, \dots, v_\ell))$ the *partial online-LOCAL run of length ℓ of G* as this is exactly the information that an online-LOCAL algorithm would know about G when deciding the output for node v_ℓ .

We denote by the pair $(\bar{G}_\ell, (w_1, \dots, w_k))$ where $\bar{G}_\ell \subseteq G_\ell$ is the connected component containing v_ℓ , and (w_1, \dots, w_k) is the maximal subsequence of (v_1, \dots, v_ℓ) of nodes that belong to \bar{G}_ℓ . In such case, k is the length of $(G_\ell, (v_1, \dots, v_\ell))[v_\ell]$. Trivially, $w_k = v_\ell$.

If $(G_\ell, (v_1, \dots, v_\ell))[v_\ell] = (G_\ell, (v_1, \dots, v_\ell))$, we say that $(G_\ell, (v_1, \dots, v_\ell))$ is a *partial one-component online-LOCAL run of length ℓ of G* . In such case, we call v_ℓ the *center* of the neighborhood.

Finally, we denote by G_ℓ^- the disjoint set of partial one-component online-LOCAL runs formed by $v_1, \dots, v_{\ell-1}$, that is, one step shorter than G_ℓ ; this is exactly the information that an online-LOCAL algorithm knows before labeling node v_ℓ when $(G_\ell, (v_1, \dots, v_\ell))$ is a partial one-component online-LOCAL run.

Definition 6.3 (Amnesiac algorithm). Let \mathcal{A} be a deterministic online-LOCAL algorithm, and let us fix the number of nodes to be n . We say that \mathcal{A} is *amnesiac* if the following condition is met:

Consider any two graphs G, H of n nodes and any adversarial orderings of the nodes (v_1, \dots, v_n) (for G) and (u_1, \dots, u_n) (for H). Fix any pair $(i, j) \in [n]^2$ of indices such that

$$\begin{aligned} (G_i, (v_1, \dots, v_i))[v_i] &= (\bar{G}_i, (v'_1, \dots, v'_k)) \text{ and} \\ (H_j, (u_1, \dots, u_j))[u_j] &= (\bar{H}_i, (u'_1, \dots, u'_k)) \end{aligned}$$

and are isomorphic (with the isomorphism being order-preserving, i.e., bringing v'_h into u'_h , for all $h \in [k]$). Then the outputs of u_i and v_j coincide.

Intuitively, \mathcal{A} is amnesiac if the output of \mathcal{A} for some node v depends only on the local connected component of the partial online local run and not anything else the algorithm has seen.

In general, an online-LOCAL algorithm uses global memory and hence it is not amnesiac.

6.2 Online to amnesiac

We are now ready to generalize the intuition we provided in Section 2.3 for turning any online-LOCAL algorithm into an amnesiac algorithm with the cost of blowing up its locality.

Lemma 6.4. *Let Π be an LCL problem with degree constraint Δ , input label set Σ_{in} and output label set Σ_{out} . Let \mathcal{A} be a deterministic online-LOCAL algorithm solving Π over a family \mathcal{F} of graphs that is closed under disjoint graph union and node/edge removal. Assume \mathcal{A} has locality $T(n)$ for an n -node graph. Then there exists an amnesiac online-LOCAL algorithm \mathcal{A}' that solves Π on an n -node graph in \mathcal{F} with locality $T(2^{O(n^3)})$.*

Proof. Let $T = T(N)$ be the locality of the following experiment; we determine the size N of the experiment graph later. Let \mathcal{G}_ℓ be the set of all partial one-component online-LOCAL runs of length ℓ for n -node graphs in \mathcal{F} (up to isomorphisms). We denote $g_\ell = |\mathcal{G}_\ell|$ and remark that $g_\ell \leq g = 2^{n^2} |\Sigma_{\text{in}}|^n$. Notice that the term 2^{n^2} takes into accounts not only all possible topologies (up to isomorphisms), but also all possible orderings of the nodes: indeed, there are at most 2^{2^n} directed graphs of n nodes where the vertices are labelled with elements of $[n]$ (which defines an ordering). We set $N_\ell = (3^{|\Sigma_{\text{out}}|} ng)^{n-\ell+1} n$.

We now adaptively construct an experiment graph H in n phases. In the first phase, we construct N_1 copies of all graphs in \mathcal{G}_1 , that is, we construct $N_1 g_1$ disjoint components describing all possible isolated neighborhoods an online-LOCAL algorithm might see. We then show the centers of these regions to \mathcal{A} and let it label them. As there are N_1 copies of each partial online local run of length 1, for each of these, say \mathcal{T} , there exists a canonical label $\sigma_{\mathcal{T}}$ that appears at least $N_1/|\Sigma_{\text{out}}|$ times, by the pigeonhole principle. We say that such online-LOCAL runs of length 1 are *good*, and ignore the rest.

In each subsequent phase ℓ , we again construct N_ℓ copies of all graphs in \mathcal{G}_ℓ , with the catch that, for each graph $(G, (v_1, \dots, v_\ell)) \in \mathcal{G}_\ell$ and for each component of G^- , we take the corresponding *good* partial one-component online-LOCAL run from a previous phase; we defer arguing why such good runs exist for now. Next we show node v_ℓ to \mathcal{A} and let it label the node. Finally, again by the pigeonhole principle, for each partial one-component online-LOCAL run \mathcal{T} (of length ℓ) we find a canonical label $\sigma_{\mathcal{T}}$ that appears at least $N_\ell/|\Sigma_{\text{out}}|$ times, and mark the corresponding runs as good.

Now our new online-LOCAL algorithm \mathcal{B} works as follows: Before even processing the first node of the input, \mathcal{B} runs the above-described experiment by simulating \mathcal{A} on H with locality T . Notice that the nodes of the experiment graph will be processed according to a *global ordering* which is *locally consistent* with that of the nodes in the single partial one-component online-LOCAL runs. We do not care about the specific global ordering of the nodes.

When processing node v , algorithm \mathcal{B} looks at partial one-component online-LOCAL run around v and finds a corresponding good and unused run from the experiment graph H . Algorithm \mathcal{B} labels v with the good label that it found in H and marks the neighborhood in H *used*.

In each step, algorithm \mathcal{B} can find such unused good neighborhoods as after the whole input has been processed, what the algorithm has seen corresponds to a partial online-LOCAL run, and in particular each component of this run corresponds to a partial one-component online-LOCAL run in H . There are at most n such component. By construction, there are at least N_n copies of each type of partial one-component runs in H . At least $N_n/|\Sigma_{\text{out}}| \gg n$ of those are good, and hence the algorithm can find good and unused neighborhoods for each of the final partial one-component online-LOCAL runs. But as each good run in H is formed by combining only good components in H , there must also be a good and unused partial one-component online-LOCAL run corresponding to each processed node v . This construction ensures that \mathcal{B} is amnesiac.

Now, by contradiction, assume that \mathcal{B} doesn't solve Π . Since Π is an LCL, there must be a node v of the input graph such that the whole output in its radius- r neighborhood, with r being the checking radius, is not admissible. However, the output of \mathcal{B} in this neighborhood is the output of \mathcal{A} from the experiment graph in a partial one-component online-LOCAL run, which must be correct by the hypothesis.

The only thing left to do is to argue that we can always find a good online-LOCAL run in the construction and to compute the final size of the experiment graph, and hence the locality of \mathcal{B} . We start with the former: Consider phase ℓ . All subsequent phases contain $\sum_{k=\ell+1}^n N_k g_k$ graphs, the simulation uses at most n graphs, and each of those may use at most n previously-used components. Hence, nodes from phase ℓ can be used at most

$$n + n \sum_{k=\ell+1}^n N_k g_k \leq n + ng \sum_{k=\ell+1}^n N_k \leq n + \frac{N_\ell}{2^{|\Sigma_{\text{out}}|}}$$

As $n \ll N_\ell$, we get that

$$\frac{N_\ell}{|\Sigma_{\text{out}}|} \geq n + n \sum_{k=\ell+1}^n N_k g_k,$$

that is the future phases won't run out of components.

Finally, the size of the experiment graph is

$$N = n \sum_{\ell=1}^n N_\ell g_\ell \leq N_0 = n(3|\Sigma_{\text{out}}|gn)^{n+1} = n2^{O(n^3)} = 2^{O(n^3)},$$

and it belongs to \mathcal{F} as \mathcal{F} is closed under disjoint graph union and node/edge removal. \square

The next lemma shows how to generalize the idea behind Lemma 6.4 all the way up to randomized online-LOCAL.

Lemma 6.5. *Let Π be an LCL problem with degree constraint Δ , input label set Σ_{in} and output label set Σ_{out} . Let \mathcal{A} be a randomized online-LOCAL algorithm solving Π over a family \mathcal{F} of graphs that is closed under disjoint graph union and node and edge removals. Assume \mathcal{A} has locality $T(n)$ and success probability $p(n) > 0$ for an n -node graph. Then there exists an amnesiac online-LOCAL algorithm \mathcal{A}' that solves Π on a graph in \mathcal{F} with n nodes with locality $T(2^{O(n^3)} + 2^{O(2^{n^2})} \cdot \log \frac{1}{p(n)})$.*

Note the locality is such that, if $p(n)$ converges to 1 fast enough, then we only have a $2^{\text{poly}(n)}$ blowup in the locality (instead of a doubly-exponential one).

Proof. The main idea is to adapt the proof of Lemma 6.4. The main challenge is that the construction of the simulation graph in the proof of Lemma 6.4 was adaptive to the output of the deterministic online-LOCAL algorithm; however, the adversary in the randomized online-LOCAL model is oblivious to the randomness of the algorithm. In fact only a $q' = 1/|\Sigma_{\text{out}}|$ fraction of components in each phase of the proof of Lemma 6.4 are good, and we cannot adaptively choose to use only those.

To combat this, we modify the experiment: instead of adaptively choosing only the good components and discarding the rest, we sample the components from the previous phases uniformly at random without replacement. We later show that the sampled component is good with probability at least $q = q'/2$. In the worst case, each component at layer N_ℓ may use up to n partial one-component online-LOCAL runs from the previous layers $N_1, \dots, N_{\ell-1}$, each of which is sampled to be good with probability only q . As the construction is n levels deep (i.e., there are n layers N_1, \dots, N_n) and the whole construction needs to work for all up to $g = 2^{n^2} |\Sigma_{\text{in}}|^n$ different partial one-component online-LOCAL runs, the total probability of finding a good label in all $N_\ell, \ell \in [n]$, is at least q^{gn^2} .

This probability is minuscule, but we can boost it by increasing the size of the construction by considering

$$k = 1 + \left\lceil q^{-2^{n^2} |\Sigma_{\text{in}}|^n n^2} \log \frac{1}{p(n)} \right\rceil$$

disjoint copies of the simulation graph independently. Letting N be the size of a single simulation graph, kN is then the size of the whole experiment. Since $N = 2^{O(n^3)}$ in Lemma 6.4, we have

$$kN = 2^{O(n^3)} + 2^{O(2^{n^2})} \log \frac{1}{p(n)}.$$

The probability that each simulation graph does not contain all the necessary good labels in all layers is at most

$$(1 - q^{gn^2})^k \leq e^{-kq^{gn^2}} = e^{-kq^{2^{n^2} |\Sigma_{\text{in}}|^n n^2}} < p(n).$$

Since the success probability of \mathcal{A} is $p(n)$, by the inclusion-exclusion principle, we get positive probability that least one simulation graph in the experiment contains all the necessary good labels and, in addition, \mathcal{A} does not fail on it.

We show that the probability of sampling a good component is indeed at least $q = 1/(2|\Sigma_{\text{out}}|)$. Consider the components in phase ℓ . For each component type, there are at least $N_\ell/|\Sigma_{\text{out}}|$ that are good. By the calculations in the proof of Lemma 6.4, future layers will use at most $N_\ell/(2|\Sigma_{\text{out}}|)$ of them. Hence, after each step, at least $1/(2|\Sigma_{\text{out}}|) = q$ fraction of the components are still good and unused, as desired.

We finish the proof by showing that there indeed exists an amnesiac algorithm \mathcal{A}' as in the claim. As argued above, our experiment succeeds with positive probability. After fixing the randomness $f : V(H) \rightarrow \{0, 1\}^{\mathbb{N}}$ of the algorithm \mathcal{A} , we obtain a uniquely defined deterministic online-LOCAL algorithm $\mathcal{A}[f]$ that behaves exactly as \mathcal{A} when given the random bit string f as input. In particular, for such f , the following holds:

1. There exists a simulation graph with the necessary number of good labels in all layers.
2. $\mathcal{A}[f]$ does not fail on the aforementioned simulation graph.

Now the amnesiac algorithm \mathcal{B} works as follows: Since there are finitely many possible simulation graphs and finitely many behaviors of $\mathcal{A}[f]$ (for all possible f), \mathcal{B} executes a preprocessing phase in which it goes over all possible simulation graphs and tries all possible deterministic online-LOCAL algorithms (according to some ordering, e.g., lexicographical in the description) until it finds the pair with the properties 1 and 2 above. We take such pair and use the labeling for that component as the base for our amnesiac algorithm. Proceeding as in the proof of Theorem 6.1, we then obtain an amnesiac algorithm that has locality

$$T \left(2^{O(n^3)} + 2^{O(2n^2)} \log \frac{1}{p(n)} \right). \quad \square$$

6.3 From online-LOCAL to SLOCAL

We first show how the SLOCAL model can nicely “partition” a rooted forest. The idea is not new and comes from [22, Section 7] where it is directly applied to the LOCAL model; we present here an adaptation to SLOCAL. The definition of rooted forest is given in Section 5. For a rooted forest $G = (V, E)$ and each node $v \in V$, we denote by G_v the subtree of G that is rooted at v . A root-to-leaf path $v_0 v_1 \dots v_k$, where $v_i \in V$ for all $i \in \{0, 1, \dots, k\}$, is a path of a rooted tree such that (v_i, v_{i-1}) is an edge for all $i \in [k]$: we say that the path *starts* at v_0 and *ends* at v_k .

Definition 6.6 ((α, β) -clustering of rooted trees). Let $G = (V, E)$ be a rooted forest. An (α, β) -clustering of G is a subset $\mathcal{L} \subseteq V$ of *leader nodes* that contains the root and is such that, for each $v \in \mathcal{L}$, the following properties are met:

1. G_v does not contain elements of \mathcal{L} at levels $1, \dots, \alpha - 1$.
2. Each maximal oriented path in G_v contains exactly one element $u \in \mathcal{L}$ such that $\text{dist}_{G_v}(u, v) \in [\alpha, \beta]$, unless the length of the maximal oriented path is at most $\beta - 1$, in which case there is at most one element of \mathcal{L} in the path (and possibly none).

A *cluster* is a maximal connected component of non-leader nodes. See Fig. 4 for an example.

We will combine many SLOCAL algorithms: it is easy to prove that the combination of two SLOCAL algorithms with localities T_1, T_2 gives an SLOCAL algorithm with locality $O(T_1 + T_2)$ [35, Lemma 2.3].

Lemma 6.7. *Let $\alpha \in \mathbb{N}_+$, and let $G = (V, E)$ be a rooted forest. There is an SLOCAL algorithm with locality $O(\alpha)$ that produces an $(\alpha - 1, 2\alpha + 1)$ -clustering of G in time at most 4α .*

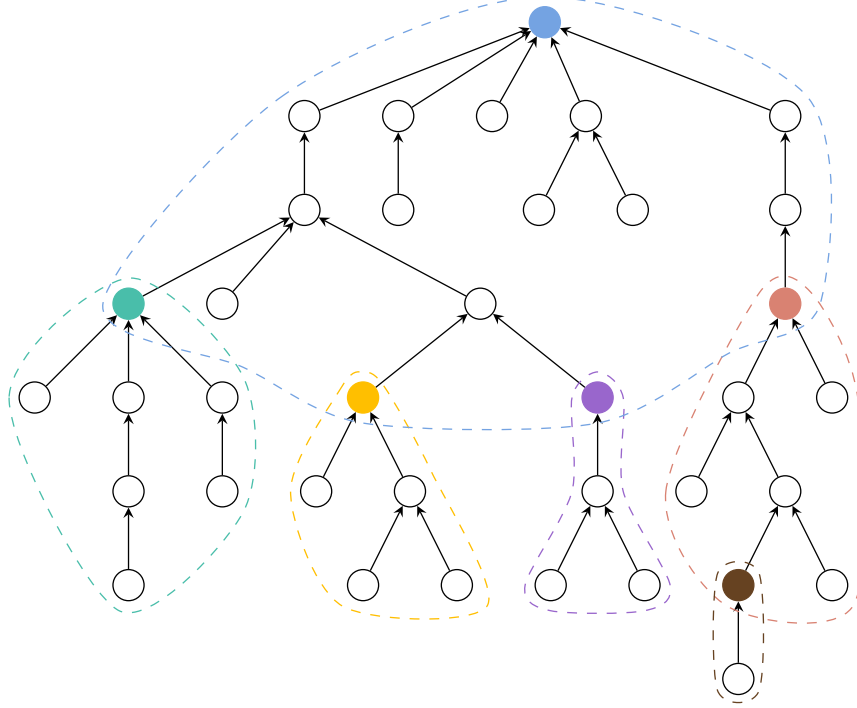


Figure 4: A $(3, 4)$ -clustering of a rooted tree. The leader nodes are colored and their clusters marked with their respective color.

Proof. We combine some SLOCAL algorithms. Consider first the following algorithm \mathcal{A}_1 : Suppose a node $v \in V$ is picked by the adversary and asked to commit to something. For each root-to-leaf path $v_0v_1 \dots v_{2\alpha-1}$ (with $v_0 = v$) that v can distinguish using locality $2\alpha - 1$, in parallel, v precommits that the node $v_{\alpha-1}$ is the *leader* of the path $v_{\alpha-1}v_\alpha \dots v_{2\alpha-1}$ unless there is another precommitment on the path within distance α from v .

Next we define an algorithm \mathcal{A}_2 that takes as input a rooted forest labelled by \mathcal{A}_1 : Each node v looks at its radius- α neighborhood and checks if there is some other (unique) node that precommitted for v in the neighborhood. If so, it becomes a leader and stores the path that is under its leadership, otherwise it does nothing.

The final algorithm \mathcal{A}_3 takes as input a rooted forest labelled by \mathcal{A}_2 and is defined as follows: Each node v becomes a leader if and only if it belongs to an oriented root-to-leaf path $v_0v_1 \dots v_{\alpha-1}$, with $v_{\alpha-1} = v$ lead by v_0 . All other nodes do not output anything, except the root, that becomes a leader.

We now prove that the SLOCAL algorithm \mathcal{A} providing the clustering is the composition of \mathcal{A}_1 , \mathcal{A}_2 , and \mathcal{A}_3 , and has locality $O(\alpha)$.

Notice that the root is a leader thanks to \mathcal{A}_3 . Property 1 in Definition 6.6 is satisfied due to \mathcal{A}_3 as well. Furthermore, for each maximal root-to-leaf path, consecutive leader nodes must be within distance at most $2\alpha + 1$ between each other, unless the path ends with a leaf at distance at most 2α from the last leader. In fact, if the distance between consecutive leaders is at least $2\alpha + 2$, or there is only one leader and the path is longer than 2α from the last leader, there would be at least one non-leader node that does not see any leader within distance α in the path, which is impossible due to how precommitments are done in \mathcal{A}_1 . \square

We can now prove Theorem 6.1, which we restate here for the reader's convenience:

Theorem 6.1. *Let Π be an LCL problem with degree constraint Δ , input label set Σ_{in} , output label set Σ_{out} , and checking radius $r > 0$. In addition, let \mathcal{A} be an online-LOCAL algorithm solving Π with locality $T(n)$*

over rooted forests. Then the following holds:

1. If \mathcal{A} is deterministic, then there exists a deterministic SLOCAL algorithm solving Π with locality $O(r) + T(2^{O(n^3)})$.
2. If \mathcal{A} is randomized and has success probability $p(n) > 0$, then there exists a deterministic SLOCAL algorithm solving Π with locality $O(r) + T(2^{O(n^3)} + 2^{O(2n^2)} \cdot \log \frac{1}{p(n)})$.

Proof. Let G denote the input graph, which is a rooted forest, and assume \mathcal{A} is a deterministic online-LOCAL algorithm. We can apply Lemma 6.4 to get an amnesiac algorithm \mathcal{A}' that solves Π with locality $T'(n) = T(2^{O(n^3)})$. We now show how to get an SLOCAL algorithm \mathcal{B} solving the problem with roughly the same locality. \mathcal{B} is the composition of four different SLOCAL algorithms $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$, each of them having locality $O(T')$:

1. We first use Lemma 6.7 with $\alpha = 10T' + 4r$, where r is the checking radius of Π , and obtain an algorithm \mathcal{B}_1 that outputs an $(\alpha - 1, 2\alpha + 1)$ -clustering of G in time 4α .
2. Now consider an algorithm \mathcal{B}_2 with locality $2\alpha + 4r$ that takes as input G as labelled by \mathcal{B}_1 and works as follows: When processing a leader node v , it collects the topology of the radius- $2r$ neighborhood of the set composed of the closest leaders v sees in each root-to-leaf path. Then, v runs \mathcal{A}' in this neighborhood and precommits a solution to the LCL for the nodes in such neighborhood. If v is the root, it also presents its own radius- $2r$ neighborhood to \mathcal{A}' and precommits a solution for the whole neighborhood. Observe that all such neighborhoods are disjoint by construction of the $(\alpha - 1, 2\alpha + 1)$ -clustering.
3. \mathcal{B}_3 also has locality $2\alpha + 4r$ and just makes all nodes whose label has been precommitted by some other node actually output the precommitted label. (The output of \mathcal{B}_2 guarantees there are no conflicts to be resolved.)
4. Finally, \mathcal{B}_4 has again locality $2\alpha + 4r$ and just brute-forces a solution in each cluster. The solution is guaranteed to exist because \mathcal{A}' was run in all disjoint neighborhoods and works correctly. In fact, \mathcal{B}_4 can just continue calling \mathcal{A}' on each cluster it sees.

The combination of $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$, and \mathcal{B}_4 yields a deterministic SLOCAL algorithm \mathcal{B} that has locality less than $13\alpha = 130T' + 52r$. (We made no attempt at optimizing the constants and they can certainly be improved upon.)

The same argument applies to a randomized online-LOCAL algorithm, the only change is that now T' is $T'(n) = T(2^{O(n^3)} + 2^{O(2n^2)} \cdot \log \frac{1}{p(n)})$. □

7 Lower bound for 3-coloring grids in randomized online-LOCAL

In this section, we will show that 3-coloring $(\sqrt{n} \times \sqrt{n})$ -grids in randomized online-LOCAL requires $\Omega(\log n)$ -locality:

Theorem 7.1. *The locality of a randomized online-LOCAL algorithm for solving 3-coloring in $(\sqrt{n} \times \sqrt{n})$ -grids is $\Omega(\log n)$.*

We will base our proof on the recent result of Chang et al. [21], where the authors show a similar lower bound for the deterministic online-LOCAL model. We will start with a brief overview of the techniques used in [21].

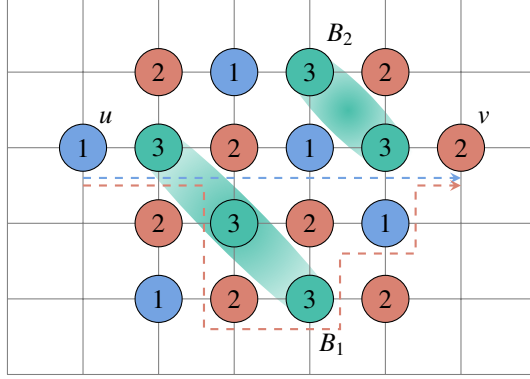


Figure 5: Example of a 3-colored grid with two boundaries B_1 and B_2 that have different parities. As one can see, $b(u, v) = 2$: No matter if we take the blue or the red path from u to v , we always cross B_1 and then B_2 .

7.1 Relevant ideas from Chang et al. [21]

The lower bound in [21] is based on the idea that any coloring of a grid G with the three colors $\{1, 2, 3\}$ can be partitioned into *regions* with colors 1 and 2 that are separated by *boundaries* of color 3. Formally, a region is a maximal connected component that is colored only with colors 1 and 2, while a boundary is a maximal connected component in G^2 that is colored only with color 3. It is crucial to observe that the boundaries of color 3 are not always compatible and that they have parities of their own (see Fig. 5). Indeed, the core idea of the proof is to constrict many incompatible boundaries of color 3 to a small space, thus requiring a view of $\Omega(\log n)$ to resolve them.

Using the same terminology as [21], we count incompatible boundaries between two points by using so-called *a- and b-values*. The *a-value* is defined as an edge weight between any two nodes and captures the change of colors 1 and 2.

Definition 7.2 (*a-value* [21]). Given a directed edge (u, v) and a 3-coloring of the nodes $c : V \mapsto \{1, 2, 3\}$, we define

$$a(u, v) = \begin{cases} c(u) - c(v), & \text{if } c(u) \neq 3 \text{ and } c(v) \neq 3 \\ 0, & \text{otherwise.} \end{cases}$$

Observe that the *a-value* of any directed 4-cycle in a grid is equal to 0. Using the *a-value*, we define the *b-value* of a path.

Definition 7.3 (*b-value* [21]). For a directed path P , its *b-value* is defined as

$$b(P) = \sum_{(u,v) \in P} a(u, v).$$

On a high level, the *b-value* of a path describes the cumulative total of incompatible boundaries along this path. Note we say ‘‘cumulative’’ because in this count boundaries of the same parity cancel each other out. Indeed, if we consider the *b-value* of a simple directed cycle in a grid, then it must be equal to 0:

Lemma 7.4 (*b-value of a cycle is zero* [21]). *Let C be a directed cycle in G . Then $b(C) = 0$.*

As an example, observe that a path that starts and ends with color 3 and otherwise is colored with colors 1 and 2 always has a *b-value* of 0 or 1. In particular, the *b-value* is 1 if the distance between the nodes of color 3 is even (and thus the boundaries are incompatible). Meanwhile, a path that goes through two incompatible regions has a total *b-value* of 2. Nevertheless, a path going through two boundaries that are compatible has a total *b-value* of 0.

Lemma 7.5 (Parity of the b -value [21]). *Let P denote any directed path of length ℓ that starts in node u and ends in node v in a grid. Then, the parity of $b(P)$ is*

$$b(P) \equiv \beta(u) + \beta(v) + \ell \pmod{2}$$

where β is an indicator variable stating whether a node is of color 3 or not:

$$\beta(u) = \begin{cases} 1, & \text{if } c(u) = 3 \\ 0, & \text{otherwise.} \end{cases}$$

Observe that the parity of the b -value of a path is determined by the colors of the endpoints of this path.

7.2 From deterministic to randomized online-LOCAL

Suppose that we are given a randomized online-LOCAL algorithm with visibility radius $T = o(\log n)$. Our goal is to prove the algorithm fails to solve the 3-coloring problem. Using Yao's minimax principle, we show how to construct an (oblivious) adversarial distribution of inputs so that any deterministic online-LOCAL algorithm fails with noticeable probability.

As in [21], the lower bound consists of two main steps:

1. First we show how to force paths to have an arbitrarily large b -value. Generally we cannot force a large b -value between the path's endpoints (or in fact between any two fixed nodes), but we do get the guarantee that it contains two nodes v_1 and v_2 where $b(v_1, v_2)$ is large. Note the location of v_1 and v_2 is completely unknown to us since we are working with an oblivious adversary. The construction is inductive: Given a procedure that generates a path with b -value $\geq k - 1$ with probability $\geq 1/2$, we show how to generate a path with b -value $\geq k$ also with probability $\geq 1/2$. (Cf. the construction in [21] with an adaptive adversary, which succeeds every time.) Since we invoke the construction for $k - 1$ a constant number of times, we are able to obtain any desired b -value of $k = o(\log n)$ with high probability. Note it is logical that we cannot do much better than this as it would contradict the existing $O(\log n)$ deterministic online-LOCAL algorithm.
2. The second step is to actually obtain a contradiction. (See Fig. 6.) First we construct a path P_1 with large b -value, say $\gg 4T$, between nodes u_s and v_t . To get the contradiction, we wish to place two nodes w_s and w_t next to u_s and v_t , respectively, so that the four node cycle has positive b -value (which is impossible due to Lemma 7.4). If we had an adaptive adversary as in [21], this would be relatively simple: Since the adaptive adversary knows the location of u_s and v_t , it just constructs an arbitrary path P_2 with the same size as P_1 , picks any two w_s and w_t that are at the same distance from each other as u_s and v_t , mirrors P_2 if needed to obtain a non-negative b -value, and then places this at minimal distance to P_1 . We show that, allowing for some failure probability, we do not need any information about u_s and v_t (i.e., nor their location nor the distance between them) in order to obtain the same contradiction.

We now proceed with the proof as outlined above. Accordingly, the first step is the following:

Lemma 7.6. *Given any $k = o(\log n)$, there is an adversarial strategy to construct a directed path of length $n^{o(1)}$ with a b -value of at least k against any online-LOCAL algorithm with locality $T = o(\log n)$. Moreover, this strategy succeeds with high probability.*

Proof. Consider the following recursive construction:

- If $k = 0$, create a single node with previously unrevealed nodes all around it (inside the visibility radius T).

- Otherwise, repeat the following steps four times in total:
 - Create four distinct paths P_1, P_2, P_3, P_4 by following the procedure for $k - 1$.
 - Toss independent fair coins $c_1, c_2, c_3 \in \{0, 1\}$.
 - Connect the P_i 's horizontally aligned and in order while placing $c_i + 1$ nodes between paths P_i and P_{i+1} .

Let Q_1, Q_2, Q_3, Q_4 be the four paths created by this procedure (each having their own four P_i 's coming from the procedure in the previous iteration $k - 1$). Next, we connect the Q_i 's horizontally aligned with each other and in arbitrary order. Note that we need to place an additional node between each pair of Q_i 's. Otherwise, we would have to have revealed the edge between the two endpoints too early to the algorithm.

Thus, for $k \geq 1$ we have 16 invocations of the procedure for $k - 1$ in total. We argue that, with probability at least $1/2$, the path yielded by this procedure contains a segment with b -value at least k . By repeating this procedure $O(\log n)$ times independently, we obtain at least one path with the desired property with high probability.

To prove that the recursive construction works, we proceed by induction. Fix $k \geq 1$ and suppose that the procedure for $k - 1$ succeeds with probability $p \geq 1/2$ at yielding a segment with b -value at least $k - 1$. Let us first consider Q_1 . Let X_i be a random variable that is 1 if this occurs for path P_i and 0 otherwise. Then we have

$$\begin{aligned} \Pr\left[\sum_{i=1}^4 X_i < 2\right] &= \Pr\left[\sum_{i=1}^4 X_i = 0\right] + \Pr\left[\sum_{i=1}^4 X_i = 1\right] \\ &= (1 - p)^4 + 4p(1 - p)^3 \\ &= (1 - p)^3(1 + 3p) \\ &\leq \frac{1}{2}. \end{aligned}$$

Hence, with probability at least $1/2$ there are at least two paths P_i and P_j , $i < j$, for which the construction succeeds.

Since we are dealing with paths, we may simplify the notation and write $b(u, v)$ for the b -value of the (unique) segment that starts at u and ends at v . Let thus u_i, v_i, u_j, v_j appear in this order in Q_1 and $|b(u_i, v_i)|, |b(u_j, v_j)| \geq k - 1$. Next we will show that, conditioned on this assumption, the probability that Q_1 contains a segment with b -value at least k is at least $1/2$. Hence, *a priori*, Q_1 contains such a segment with probability at least $1/4$. Since all Q_i 's are constructed the in the same way and independently of one another, the probability that at least one of the Q_i 's contains such a segment is at least $1 - (1 - 1/4)^4 > 1 - 1/e > 1/2$.

Let us write $\sigma(x)$ for the sign function of x (i.e., $\sigma(x) = 1$ if $x > 0$, $\sigma(x) = -1$ if $x < 0$, and $\sigma(0) = 0$). To see why the above holds for Q_1 , consider the two following cases:

Case 1: $\sigma(b(u_i, v_i)) = \sigma(b(u_j, v_j))$. Using Lemma 7.5, we have that $b(v_i, u_j) \equiv \beta(v_i) + \beta(u_j) + c_i + \dots + c_{j-1} \pmod{2}$. Since $\beta(v_i)$ and $\beta(u_j)$ are fixed, and the coin tosses are independent, we have $|b(v_i, u_j)| \not\equiv k - 1 \pmod{2}$ with probability $1/2$. Assuming this holds, we have either $|b(v_i, u_j)| \geq k$, in which case we are done, or $|b(v_i, u_j)| \leq k - 2$. Since $b(u_i, v_i)$ and $b(u_j, v_j)$ have the same sign, we have

$$\begin{aligned} |b(u_i, v_j)| &= |b(u_i, v_i) + b(v_i, u_j) + b(u_j, v_j)| \\ &\geq |b(u_i, v_i) + b(u_j, v_j)| - |b(v_i, u_j)| \\ &\geq 2(k - 1) - (k - 2) \\ &= k. \end{aligned}$$

Case 2: $\sigma(b(u_i, v_i)) \neq \sigma(b(u_j, v_j))$. Arguing by using Lemma 7.5 as before, we obtain that $|b(v_i, u_j)| \not\equiv 0 \pmod{2}$ holds with probability $1/2$. Assuming this is the case, we have thus $|b(v_i, u_j)| \geq 1$. Without restriction, let $\sigma(b(u_i, v_i)) = \sigma(b(v_i, u_j))$. Then

$$|b(u_i, u_j)| = |b(u_i, v_i) + b(v_i, u_j)| \geq k - 1 + 1 = k.$$

Finally, let us confirm that the construction fits into the $(\sqrt{n} \times \sqrt{n})$ -grid. We only reveal at most $2T + 1 = o(\sqrt{n})$ in a column, so we need to only consider nodes along a row. The initial path contains $m_0 = 2T + 1$ visible nodes and in the i -th recursive step we have $m_i \leq 16p_{i-1} + 27$ visible nodes in total for $i \geq 1$. (Inside each Q_i we need at most $2(4 - 1) = 6$ additional nodes to join the four P_i 's, and to join the four Q_i 's we need an additional 3 nodes.) Solving the recursion, in the k -th step we have thus

$$m_k = \frac{1}{5} \left(2^{4k+1} (5T + 7) - 9 \right) = n^{o(1)}$$

visible nodes along the path since $k, T = o(\log n)$. □

We now illustrate the idea for getting the contradiction previously described. (See Fig. 6.) Invoking Lemma 7.6, we obtain a path P_1 with a b -value of $4T + 4$ between two nodes u_s and v_t (whose positions are unknown to us). Letting L be the length of P_1 , we (arbitrarily) create a second path P_2 of length $10L$ and then randomly choose how to align P_1 and P_2 . More specifically, letting u be the first node in P_1 , we choose some node w of P_2 uniformly at random and align u and w ; then we mirror P_2 with probability $1/2$ and reveal it at distance $2T + 2$ to P_1 (which is consistent with all previously revealed nodes).

The reason why this works is the following: Since P_1 has length L and P_2 length $10L$, with at least $4/5$ we align the paths so that each node in P_1 has a matching node underneath it in P_2 . Let w_s and w_t be the nodes matching u_s and v_t , respectively. Then because we mirror P_2 with probability $1/2$, we get that $b(w_s, \dots, w_t)$ is at least zero in expectation (conditioned on having properly aligned the two paths). Hence, with probability at least $2/5$ we obtain a cycle $(u_s, \dots, v_t, \dots, w_t, \dots, w_s, \dots, u_s)$ where $b(u_s, \dots, v_t) > 4T + 4$, $b(w_t, \dots, w_s) \geq 0$, and $b(v_t, \dots, w_t), b(w_s, \dots, u_s) \geq -2T - 2$ (due to the distance between the two paths).

Proof of Theorem 7.1. Observe that, using Lemma 7.6, we can construct a path $P_1 = (u_0, \dots, u_L)$ of length $L = n^{o(1)}$ where some segment (u_s, \dots, u_t) of P_1 has a b -value of $k > 4T + 4$. Next we construct a path $P_2 = (v_0, \dots, v_{10L})$ of length $10L$ that will be placed below P_1 . We assume that the points of the path are revealed to the algorithm in some predefined order.

The position and orientation of P_2 are chosen as follows:

1. Choose a node $v_r \in [L, 9L]$ of P_2 uniformly at random. This node is placed below the node u_0 of P_1 at $2T + 2$ distance from it.
2. Throw a fair coin $m \in \{0, 1\}$. If $m = 1$, mirror P_2 along the vertical axis that goes through u_0 and v_r .

The nodes of P_2 are revealed to the algorithm in the same predefined and possibly mirrored order. See Fig. 6 for an example.

We next prove that we obtain the desired lower bound. First notice that, since we pick $v_r \in [L, 9L]$, every node in P_1 has a counterpart in P_2 whether we mirror P_2 or not. Let (w_s, \dots, w_t) denote the segment matched to (u_s, \dots, u_t) . Consider the case where either of the following is true:

- $m = 0$ and $r \in [L, 9L - s - t]$
- $m = 1$ and $r \in [L + s + t, 9L]$

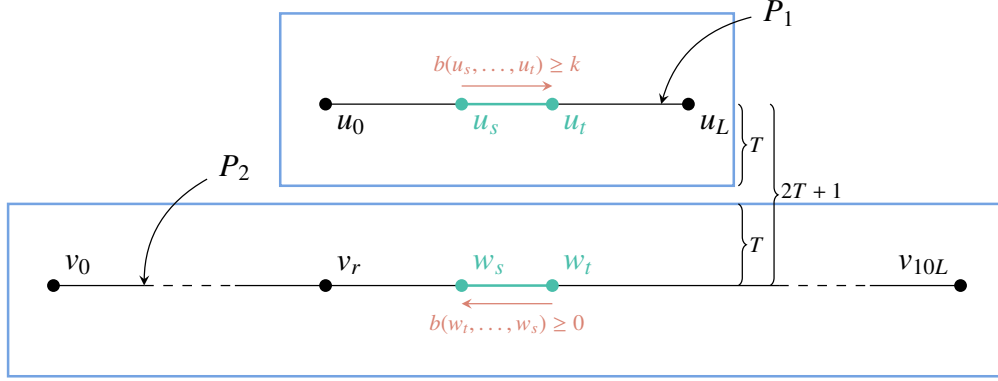


Figure 6: How to turn paths with large b -value into a contradiction. Here the blue area includes the nodes revealed so far around the path segments (u_0, \dots, u_L) and the corresponding part of P_2 underneath it. The green segments are the two segments we consider in the proof. A cycle going through both paths leads to a contradiction.

Denoting by E the event in which this occurs, note we have

$$\Pr[E] = \frac{1}{2} \Pr[r \in [L, 9L - s - t]] + \frac{1}{2} \Pr[r \in [L + s + t, 9L]] = \frac{8L - s - t + 1}{9L + 1} \geq \frac{6L + 1}{9L + 1} > \frac{2}{3}.$$

Now conditioned on E , notice that every segment (v_x, \dots, v_y) where $x \in [L + s, 9L - s]$ and $y \in [L + t, 9L - t]$ has equal probability of being matched with (u_s, \dots, u_t) either in the same direction (i.e., $w_s = v_x$ and $u_t = v_y$) or reversed (i.e., $w_t = v_y$ and $w_s = v_x$). Hence,

$$\Pr[b(w_t, \dots, w_s) \geq 0 \mid E] \geq \frac{1}{2}.$$

(In fact, the probability is exactly $1/2$ if $b(w_t, \dots, w_s) > 0$ and 1 if $b(w_t, \dots, w_s) = 0$.) Thus, the probability that $b(w_t, \dots, w_s) \geq 0$ is $> (2/3) \cdot (1/2) = 1/3$.

From here on, we proceed as in [21]. By Lemma 7.4, the cycle $(u_s, \dots, u_t, \dots, w_t, \dots, w_s, \dots, u_s)$ must have a b -value of 0 . However, recall that the b -value of a path is bounded by its length by definition. In our case, $-2T - 2 < b(u_t, \dots, w_t) < 2T + 2$, $-2T - 2 < b(w_s, \dots, u_s) < 2T + 2$, $b(u_s, \dots, u_t) \geq k$, and $b(w_t, \dots, w_s) \geq 0$. In order for the b -value of the cycle to be 0 , we would need to have $2(2T + 2) \geq k$, which is a contradiction. Since this occurs with noticeable probability (i.e., $> 1/3$), the claim follows. \square

8 Dynamic-LOCAL derandomizes LOCAL and breaks symmetry

This section presents a derandomization result for the dynamic-LOCAL model: we show that not only dynamic-LOCAL can simulate randomized LOCAL with no overhead in the locality, but actually it brings the complexity class $O(\log^* n)$ in randomized LOCAL down to $O(1)$.

We make use of the method of conditional expectations from [36] to derandomize the model of dynamic-LOCAL computation. We note that while the following theorems are proven for the class of all graphs, they can be extended to any subclass of graphs.

Theorem 8.1. *Let Π be an LCL problem, and let \mathcal{A} be a randomized LOCAL algorithm solving Π with locality $T(n)$ and probability $p > 1 - 1/n$. Then there exists a deterministic dynamic-LOCAL algorithm \mathcal{A}' solving Π with locality $O(T(n))$.*

Proof. The randomized LOCAL algorithm \mathcal{A} can be viewed as a deterministic LOCAL algorithm which is given i.i.d. random variables R_v locally for each node v . We show that there exists a dynamic-LOCAL algorithm with locality $O(T(n))$ that assigns fixed values ρ_v for each R_v such that when \mathcal{A} is run on these values as R_v , it produces a correct output for problem Π . Such dynamic-LOCAL clearly implies an $O(T(n))$ -locality dynamic-LOCAL algorithm for problem Π as one can compose it with \mathcal{A} . The proof uses a technique similar to that of derandomizing LOCAL algorithms in SLOCAL [36].

We use the method of conditional expectation. Let F_v be a flag that is 1 when the output around node v is invalid, and 0 otherwise. Note that F_v depends only on the radius- r neighborhood of node v . Let $F = \sum_{v \in V} F_v$. As Π is an LCL problem, $F = 0$ iff the labeling is correct for the whole graph, and otherwise $F \geq 1$. By definition, $\mathbb{E}[F] = \sum_{v \in V} \mathbb{E}[F_v] \leq n(1 - p) < 1$.

Let $\sigma = e_1, e_2, \dots, e_m$ be the order of edges in which the adversary dynamically modifies the graph, and let G_i be the graph after modification i . Note that any edge can appear multiple times in the sequence if the adversary also removes edges.

Now algorithm \mathcal{A}' sets the randomness as follows: In step i , algorithm \mathcal{A}' first clears ρ_u for all nodes u with $\text{dist}(u, e_i) \leq T(n) + r$. Then

$$\mathbb{E}[F \mid \bigwedge_{u, \text{dist}(u, e_i) > T(n) + r} R_u = \rho_u] < 1$$

as

$$\mathbb{E}[F \mid \bigwedge_u R_u = \rho_u] < 1$$

in the previous step. Hence, there exists some concrete values for ρ_u for $\text{dist}(u, e_i) \leq T(n) + r$ such that $\mathbb{E}[F \mid \bigwedge_u R_u = \rho_u] < 1$. Algorithm \mathcal{A}' sets all ρ_u to such values. Finally, algorithm \mathcal{A}' simulates \mathcal{A} for all nodes within radius $T(n) + r$ from e_i with randomness fixed at ρ .

This produces a correct labeling for all nodes as $\mathbb{E}[F \mid \bigwedge_u R_u = \rho_u] < 1$, and hence

$$\mathbb{E}[F \mid \bigwedge_u R_u = \rho_u] = 0. \quad \square$$

Theorem 8.2. *Let Π be an LCL problem, and let there exist a randomized LOCAL algorithm solving it with locality $O(\log^* n)$. Then there exists a dynamic-LOCAL algorithm solving Π with locality $O(1)$.*

Proof. Using [20], a randomized LOCAL algorithm for solving Π with locality $O(\log^* n)$ implies the existence of a deterministic LOCAL algorithm solving Π with locality $O(\log^* n)$. Moreover, the deterministic LOCAL algorithm can be normalized into an algorithm that first finds a distance- k coloring with $O(\Delta^k)$ colors for some constant k , and then applies an $O(k)$ -locality LOCAL algorithm \mathcal{A} on this coloring.

The dynamic-LOCAL algorithm for solving Π maintains a distance- k coloring with $O(\Delta^k)$ colors on the graph, and then applies \mathcal{A} on these colors. When the dynamic-LOCAL algorithm encounters an update, it first clears the colors of all nodes within distance k from the update, and greedily assigns new colors for those nodes. The greedy assignment always succeeds since each node can have at most Δ^k neighbors within distance k hence they can use at most Δ^k of the available colors. Finally, the algorithm applies \mathcal{A} on all nodes of the graph. As \mathcal{A} is a deterministic LOCAL algorithm with locality $O(k)$, and only the colors of nodes within distance k of the update have changed, we have that only the output of nodes within distance $O(k) = O(1)$ from the update change. \square

9 Randomized online-LOCAL with an adaptive adversary

This section aims at showing that randomized online-LOCAL with an adaptive adversary is as strong as its deterministic counterpart. The proof is very similar to the proof of Theorem 8.1 and uses, again, the method

of conditional expectations. Also, as in the previous section, the results can be extended to any subclass of graphs.

Theorem 9.1. *Let Π be an LCL problem, and let \mathcal{A} be an online-LOCAL algorithm with an adaptive adversary solving Π with locality $T(n)$ and probability $p > 1 - 1/n$. Then there exists a deterministic online-LOCAL algorithm \mathcal{A}' solving Π with locality $T(n)$.*

Proof. Let \mathcal{A} be an algorithm in the online-LOCAL model with an adaptive adversary and with round complexity $T(n)$. We construct a deterministic online-LOCAL algorithm \mathcal{A}' with the same round complexity $T(n)$ as follows: The algorithm \mathcal{A}' simulates \mathcal{A} with certain fixed random-bit-strings. In particular, whenever new nodes arrive at the input, the algorithm \mathcal{A}' fixes the random bits in all newly arrived nodes in the way we describe below.

We view a run of an online-LOCAL algorithm with an adaptive adversary as a game between an algorithm \mathcal{A} and an adversary. In each step of that game, first the adversary chooses a set of nodes that arrive at the input (one special node and its neighborhood). Then the algorithm \mathcal{A} samples the random bits of the newly arrived nodes and fixes the label of the arrived node. After i steps of this game, we use \mathcal{F} to denote the event that the adversary wins the game, that is algorithm \mathcal{A} produces an invalid output according to problem Π . By definition, we have $\Pr[\mathcal{F}] \leq 1/n$. Here and from now on, the probability expressions assume that the adversary maximizes the probability of \mathcal{F} happening.

Our algorithm \mathcal{A}' in each i -th step fixes the random bits of the newly arrived nodes in a way that guarantees that

$$\Pr[\mathcal{F}|B_1 = b_1, \dots, B_i = b_i] \leq \Pr[\mathcal{F}|B_1 = b_1, \dots, B_{i-1} = b_{i-1}].$$

Here, $B_j = b_j$ stands for fixing the values of random bits of nodes that arrived in the j -th step to b_j . We note that this can always be done since, by definition, we have

$$\Pr[\mathcal{F}|B_1 = b_1, \dots, B_{i-1} = b_{i-1}] = \mathbb{E}_{B_i}[\Pr[\mathcal{F}|B_1 = b_1, \dots, B_i = b_i]].$$

Moreover, the online-LOCAL algorithm importantly both knows all the past arrived nodes and their randomness, and can also consider all the different choices of the adversary in the future, and hence can compute the above conditional probabilities.

At the end of the simulation of \mathcal{A} , we have $\Pr[\mathcal{F}|B_1 = b_1, \dots, B_n = b_n] < 1$, and hence the bad event \mathcal{F} does not occur. This implies that the algorithm \mathcal{A}' is always correct. \square

Remark 9.2. This proof assumes that we have white-box access to the randomized online-LOCAL algorithm and that it uses finitely many random bits. If either of these assumptions does not hold, we can use a more inefficient derandomization along the lines of [29]. Here, we just observe that an online-LOCAL algorithm working against an adaptive adversary implies the existence of a function f that maps all possible input sequences on all possible n -node graphs that can be produced by the adversary to a valid output of the problem Π . The deterministic online-LOCAL algorithm finds such a function f at the beginning of its execution and produces the output according to f .

10 LCL problems in paths and cycles in randomized online-LOCAL

In this section, we show that the complexity of LCLs in paths and cycles is either $O(1)$ or $\Theta(n)$ in randomized online-LOCAL (with an oblivious adversary).

Theorem 10.1. *Let Π be an LCL problem on paths and cycles (possibly with inputs). If the locality of Π is T in the randomized online-LOCAL model, then its locality is $O(T + \log^* n)$ in the LOCAL model.*

This theorem is a simple corollary of the following two lemmas:

Lemma 10.2. *Let Π be an LCL problem on paths and cycles (possibly with inputs), and let \mathcal{A} be a randomized online-LOCAL algorithm solving Π with locality $o(n)$. Then there exists a randomized online-LOCAL algorithm \mathcal{A}' solving Π with locality $O(1)$.*

Lemma 10.3. *Let Π be an LCL problem on paths and cycles (possibly with inputs), and let \mathcal{A} be a randomized online-LOCAL algorithm solving Π with locality $O(1)$. Then there exists a LOCAL algorithm \mathcal{A}' solving Π with locality $O(\log^* n)$.*

By previous results [2], it is known that, in the case of paths and cycles, any (deterministic) online-LOCAL algorithm with sublinear locality can be sped up to an online-LOCAL algorithm with constant locality. We show how to extend [2, Lemma 5.5]. We start by constructing a large virtual graph P' such that, when the original algorithm runs on the virtual graph P' , the labeling produced by the algorithm is locally compatible with the labeling in the original graph P .

Proof of Lemma 10.2. Let Π be an LCL problem in paths or cycles with checking-radius r , let P be a path or a cycle with n nodes, and let \mathcal{A} be a randomized online-LOCAL algorithm solving Π with locality $T(n) = o(n)$. In the proof of [2, Lemma 5.5], the authors use three phases to speed up (deterministic) online-LOCAL with sublinear locality to only constant locality. We use the same strategy to construct a randomized online-LOCAL algorithm \mathcal{A}' for solving Π with constant locality. The phases for constructing randomized online-LOCAL algorithm \mathcal{A}' are as follows:

1. In the first phase, the algorithm deterministically creates an (α, α) -ruling set R for the path P , mirroring the approach outlined in the proof of [2, Lemma 5.5].
2. In the second phase, the algorithm constructs a larger virtual path P' with N nodes and simulates algorithm \mathcal{A} on P' in the neighborhoods of nodes in R . Intuitively this path P' is constructed based on the pumping-lemma-style argument on LCL problems presented by Chang and Pettie [18], and it only relies on (the definition of) problem Π , not the algorithm \mathcal{A} . During this simulation, the algorithm encounters a failure only if \mathcal{A} fails within specific neighborhoods of P' . Consequently, the algorithm's success in this phase aligns with the success rate of \mathcal{A} , ensuring a high probability of success.
3. The third phase is also as in [2, Lemma 5.5]: The algorithm extends the fixed labels around R to the entire path P by applying brute force for nodes outside the neighborhood of R . Note that this extending is feasible because of the way the path P' is created and the pumping-lemma-style argument as discussed in [2].

In the end, we compose these three phases together to obtain the randomized online-LOCAL algorithm \mathcal{A}' for solving Π with constant locality. □

Proof of Lemma 10.3. This proof closely follows the argument presented in [2, Lemma 5.6]. Let Π denote an LCL problem with a constant checking-radius r , and suppose \mathcal{A} is a randomized online-LOCAL algorithm solving Π with constant locality T . Define $\beta = T + r + 1$. As in [2, Lemma 5.6], we consider an input-labeled graph P formed by many copies of all feasible input neighborhoods with a radius of β . We can depict P as a collection of disjoint path fragments that we later connect to each other and create a long path with. Each of these path fragments has size $2\beta + 1$, so the node in the center of each segment (i.e., the $(\beta + 1)$ -th node in the segment) has the same view (up to radius β) as in the final path. Let V be the set of central nodes within these fragments.

We apply \mathcal{A} to each node within the radius- r neighborhood of the nodes in V following an arbitrary order and then terminate. These nodes have the same view as in the final path because their distance to the

endpoints is at most $\beta - r = T + 1$. We iterate the process multiple times, yielding a distribution of output labels around the central nodes. Given that the size of P is constant, there exists an output labeling L of the nodes occurring with probability $p = \Omega(1)$. If needed, we can augment P with (constantly many) additional nodes such that $|P| > 1/p$.

We set this labeling L as the deterministic output for graph P and proceed similarly to the deterministic case by constructing the canonical labeling f from L . Then, we use the function f to construct a LOCAL algorithm with $O(\log^* n)$ locality following the same steps as in the proof of [2, Lemma 5.6]. It is important to highlight that it is feasible to fill any gap of sufficient length between parts labeled with the canonical labeling: If the algorithm \mathcal{A} would never produce a valid labeling for this gap, then the algorithm would fail to label P with probability at least p and, since $p \geq 1/|P|$, the algorithm \mathcal{A} would not succeed with high probability. \square

Acknowledgments

This work was supported in part by the Research Council of Finland, Grant 333837. Xavier Coiteux-Roy is supported by the Swiss National Science Foundation. Francesco d'Amore is supported by MUR FARE 2020 - Project PARECoDi CUP J43C22000970001. Darya Melnyk is supported by the European Research Council (ERC), grant agreement No. 864228 (AdjustNet), Horizon 2020, 2020-2025. Augusto Modanese and Shreyas Pai are supported by the Helsinki Institute for Information Technology (HIIT). Marc-Olivier Renou acknowledges funding by INRIA through the Action Exploratoire project DEPARTURE, and by the ANR for the JCJC grant LINKS (ANR-23-CE47-0003).

Errata. An earlier version of this manuscript (arXiv:2403.01903v1, posted on 2024-03-04) had a mistake: we claimed a result analogous to Theorem 1.5 for *unrooted* trees (which would also have implications on the complexity of the sinkless orientation problem). However, the proof had a missing step, and we do not know if the proof can be fixed.

References

- [1] Jon Aaronson, David Gilat, Michael Keane, and Vincent de Valk. An algebraic construction of a class of one-dependent processes. *The Annals of Probability*, 17(1):128–143, 1989.
- [2] Amirreza Akbari, Navid Eslami, Henrik Lievonen, Darya Melnyk, Joonas Särkijärvi, and Jukka Suomela. Locality in online, dynamic, sequential, and distributed graph algorithms. In Kousha Etessami, Uriel Feige, and Gabriele Puppis, editors, *50th International Colloquium on Automata, Languages, and Programming, ICALP 2023, July 10-14, 2023, Paderborn, Germany*, volume 261 of *LIPICs*, pages 10:1–10:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICs.ICALP.2023.10.
- [3] Heger Arfaoui and Pierre Fraigniaud. What can be computed without communications? *SIGACT News*, 45(3):82–104, 2014. doi:10.1145/2670418.2670440.
- [4] Alkida Balliu, Juho Hirvonen, Janne H. Korhonen, Tuomo Lempinen, Dennis Olivetti, and Jukka Suomela. New classes of distributed time complexity. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1307–1318. ACM, 2018. doi:10.1145/3188745.3188860.

- [5] Alkida Balliu, Sebastian Brandt, Yi-Jun Chang, Dennis Olivetti, Mikaël Rabie, and Jukka Suomela. The distributed complexity of locally checkable problems on paths is decidable. In Peter Robinson and Faith Ellen, editors, *Proceedings of the 2019 ACM Symposium on Principles of Distributed Computing, PODC 2019, Toronto, ON, Canada, July 29 - August 2, 2019*, pages 262–271. ACM, 2019. doi:10.1145/3293611.3331606.
- [6] Alkida Balliu, Sebastian Brandt, Dennis Olivetti, and Jukka Suomela. How much does randomness help with locally checkable problems? In Yuval Emek and Christian Cachin, editors, *PODC '20: ACM Symposium on Principles of Distributed Computing, Virtual Event, Italy, August 3-7, 2020*, pages 299–308. ACM, 2020. doi:10.1145/3382734.3405715.
- [7] Alkida Balliu, Sebastian Brandt, Juho Hirvonen, Dennis Olivetti, Mikaël Rabie, and Jukka Suomela. Lower bounds for maximal matchings and maximal independent sets. *Journal of the ACM*, 68(5): 39:1–39:30, 2021. doi:10.1145/3461458.
- [8] Alkida Balliu, Sebastian Brandt, Dennis Olivetti, and Jukka Suomela. Almost global problems in the LOCAL model. *Distributed Computing*, 34(4):259–281, 2021. doi:10.1007/S00446-020-00375-2.
- [9] Alkida Balliu, Sebastian Brandt, Yi-Jun Chang, Dennis Olivetti, Jan Studený, and Jukka Suomela. Efficient classification of locally checkable problems in regular trees. In Christian Scheideler, editor, *36th International Symposium on Distributed Computing, DISC 2022, October 25-27, 2022, Augusta, Georgia, USA*, volume 246 of *LIPICs*, pages 8:1–8:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICs.DISC.2022.8.
- [10] Alkida Balliu, Juho Hirvonen, Darya Melnyk, Dennis Olivetti, Joel Rybicki, and Jukka Suomela. Local mending. In Merav Parter, editor, *Structural Information and Communication Complexity - 29th International Colloquium, SIROCCO 2022, Paderborn, Germany, June 27-29, 2022, Proceedings*, volume 13298 of *Lecture Notes in Computer Science*, pages 1–20. Springer, 2022. doi:10.1007/978-3-031-09993-9_1.
- [11] Alkida Balliu, Sebastian Brandt, Yi-Jun Chang, Dennis Olivetti, Jan Studený, Jukka Suomela, and Aleksandr Tereshchenko. Locally checkable problems in rooted trees. *Distributed Computing*, 36(3): 277–311, 2023. doi:10.1007/S00446-022-00435-9.
- [12] Alkida Balliu, Janne H. Korhonen, Fabian Kuhn, Henrik Lievonen, Dennis Olivetti, Shreyas Pai, Ami Paz, Joel Rybicki, Stefan Schmid, Jan Studený, Jukka Suomela, and Jara Uitto. Sinkless orientation made simple. In Telikepalli Kavitha and Kurt Mehlhorn, editors, *2023 Symposium on Simplicity in Algorithms, SOSA 2023, Florence, Italy, January 23-25, 2023*, pages 175–191. SIAM, 2023. doi:10.1137/1.9781611977585.CH17.
- [13] Salman Beigi and Marc-Olivier Renou. Covariance decomposition as a universal limit on correlations in networks. *IEEE Transactions on Information Theory*, 68(1):384–394, 2021.
- [14] Max Born. Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik*, 38(11-12):803–827, 1926. doi:10.1007/bf01397184.
- [15] Sebastian Brandt, Orr Fischer, Juho Hirvonen, Barbara Keller, Tuomo Lempäinen, Joel Rybicki, Jukka Suomela, and Jara Uitto. A lower bound for the distributed Lovász local lemma. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 479–488. ACM, 2016. doi:10.1145/2897518.2897570.

- [16] Sebastian Brandt, Juho Hirvonen, Janne H. Korhonen, Tuomo Lempiäinen, Patric R. J. Östergård, Christopher Purcell, Joel Rybicki, Jukka Suomela, and Przemyslaw Uznanski. LCL problems on grids. In Elad Michael Schiller and Alexander A. Schwarzmann, editors, *Proceedings of the ACM Symposium on Principles of Distributed Computing, PODC 2017, Washington, DC, USA, July 25-27, 2017*, pages 101–110. ACM, 2017. doi:10.1145/3087801.3087833.
- [17] Keren Censor-Hillel, Orr Fischer, François Le Gall, Dean Leitersdorf, and Rotem Oshman. Quantum distributed algorithms for detection of cliques. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPICs*, pages 35:1–35:25. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPICs.ITCS.2022.35.
- [18] Yi-Jun Chang and Seth Pettie. A time hierarchy theorem for the LOCAL model. *SIAM Journal on Computing*, 48(1):33–69, 2019. doi:10.1137/17M1157957.
- [19] Yi-Jun Chang, Tsvi Kopelowitz, and Seth Pettie. An exponential separation between randomized and deterministic complexity in the LOCAL model. In Irit Dinur, editor, *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 615–624. IEEE Computer Society, 2016. doi:10.1109/FOCS.2016.72.
- [20] Yi-Jun Chang, Tsvi Kopelowitz, and Seth Pettie. An exponential separation between randomized and deterministic complexity in the LOCAL model. *SIAM Journal on Computing*, 48(1):122–143, 2019. doi:10.1137/17M117537.
- [21] Yi-Jun Chang, Gopinath Mishra, Hung Thuan Nguyen, Mingyang Yang, and Yu-Cheng Yeh. A tight lower bound for 3-coloring grids in the online-LOCAL model. *CoRR*, abs/2312.01384, 2023. doi:10.48550/ARXIV.2312.01384.
- [22] Yi-Jun Chang, Jan Studený, and Jukka Suomela. Distributed graph problems through an automata-theoretic lens. *Theoretical Computer Science*, 951:113710, 2023. doi:10.1016/J.TCS.2023.113710.
- [23] Giulio Chiribella and Robert W. Spekkens, editors. *Quantum Theory: Informational Foundations and Foils*. Springer Netherlands, 2016. doi:10.1007/978-94-017-7303-4.
- [24] Xavier Coiteux-Roy, Elie Wolfe, and Marc-Olivier Renou. Any physical theory of nature must be boundlessly multipartite nonlocal. *Physical Review A*, 104(5):052207, 2021.
- [25] Xavier Coiteux-Roy, Elie Wolfe, and Marc-Olivier Renou. No bipartite-nonlocal causal theory can explain nature’s correlations. *Physical Review Letters*, 127(20):200401, 2021. doi:10.1103/physrevlett.127.200401.
- [26] Xavier Coiteux-Roy, Francesco d’Amore, Rishikesh Gajjala, Fabian Kuhn, François Le Gall, Henrik Lievonen, Augusto Modanese, Marc-Olivier Renou, Gustav Schmid, and Jukka Suomela. No distributed quantum advantage for approximate graph coloring. In *Proc. 56th Annual ACM Symposium on Theory of Computing (STOC 2024)*. ACM Press, 2024.
- [27] Xavier Coiteux-Roy, Owidiusz Makuta, Fionnuala Curran, Remigiusz Augusiak, and Marc-Olivier Renou. The genuinely multipartite nonlocality of graph states is model-dependent, 2024. In preparation.
- [28] Richard Cole and Uzi Vishkin. Deterministic coin tossing with applications to optimal parallel list ranking. *Information and Control*, 70(1):32–53, 1986. doi:10.1016/S0019-9958(86)80023-7.

- [29] Sameep Dahal, Francesco D’Amore, Henrik Lievonon, Timothé Picavet, and Jukka Suomela. Brief announcement: Distributed derandomization revisited. In Rotem Oshman, editor, *37th International Symposium on Distributed Computing, DISC 2023, October 10-12, 2023, L’Aquila, Italy*, volume 281 of *LIPICs*, pages 40:1–40:5. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICs.DISC.2023.40.
- [30] Giacomo Mauro D’Ariano, Giulio Chiribella, and Paolo Perinotti. *Quantum Theory from First Principles: An Informational Approach*. Cambridge University Press, 2016. doi:10.1017/9781107338340.
- [31] Michael Elkin, Hartmut Klauck, Danupon Nanongkai, and Gopal Pandurangan. Can quantum communication speed up distributed computation? In Magnús M. Halldórsson and Shlomi Dolev, editors, *ACM Symposium on Principles of Distributed Computing, PODC ’14, Paris, France, July 15-18, 2014*, pages 166–175. ACM, 2014. doi:10.1145/2611462.2611488.
- [32] Manuela Fischer and Mohsen Ghaffari. Sublogarithmic distributed algorithms for Lovász local lemma, and the complexity hierarchy. In Andréa W. Richa, editor, *31st International Symposium on Distributed Computing, DISC 2017, October 16-20, 2017, Vienna, Austria*, volume 91 of *LIPICs*, pages 18:1–18:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPICs.DISC.2017.18.
- [33] Cyril Gavoille, Adrian Kosowski, and Marcin Markiewicz. What can be observed locally? In Idit Keidar, editor, *Distributed Computing, 23rd International Symposium, DISC 2009, Elche, Spain, September 23-25, 2009. Proceedings*, volume 5805 of *Lecture Notes in Computer Science*, pages 243–257. Springer, 2009. doi:10.1007/978-3-642-04355-0_26.
- [34] Mohsen Ghaffari and Hsin-Hao Su. Distributed degree splitting, edge coloring, and orientations. In Philip N. Klein, editor, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 2505–2523. SIAM, 2017. doi:10.1137/1.9781611974782.166.
- [35] Mohsen Ghaffari, Fabian Kuhn, and Yannic Maus. On the complexity of local distributed graph problems. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 784–797. ACM, 2017. doi:10.1145/3055399.3055471.
- [36] Mohsen Ghaffari, David G. Harris, and Fabian Kuhn. On derandomizing local distributed algorithms. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 662–673. IEEE Computer Society, 2018. doi:10.1109/FOCS.2018.00069.
- [37] Mohsen Ghaffari, Juho Hirvonen, Fabian Kuhn, Yannic Maus, Jukka Suomela, and Jara Uitto. Improved distributed degree splitting and edge coloring. *Distributed Computing*, 33(3-4):293–310, 2020. doi:10.1007/S00446-018-00346-8.
- [38] Nicolas Gisin, Jean-Daniel Bancal, Yu Cai, Patrick Remy, Armin Tavakoli, Emmanuel Zambrini Cruzeiro, Sandu Popescu, and Nicolas Brunner. Constraints on nonlocality in networks from no-signaling and independence. *Nature Communications*, 11(1):2378, 2020. doi:10.1038/s41467-020-16137-4.
- [39] Andrew V. Goldberg, Serge A. Plotkin, and Gregory E. Shannon. Parallel symmetry-breaking in sparse graphs. *SIAM Journal on Discrete Mathematics*, 1(4):434–446, 1988. doi:10.1137/0401044.

- [40] Alexander E. Holroyd. Symmetrization for finitely dependent colouring. *CoRR*, abs/2305.13980, 2023. doi:10.48550/arXiv.2305.13980.
- [41] Alexander E. Holroyd and Thomas M. Liggett. Finitely dependent coloring. *Forum of Mathematics, Pi*, 4:e9, 2016. doi:10.1017/fmp.2016.7.
- [42] Alexander E. Holroyd, Oded Schramm, and David B. Wilson. Finitary coloring. *The Annals of Probability*, 45(5):2867–2898, 2017. doi:10.1214/16-aop1127.
- [43] Alexander E. Holroyd, Tom Hutchcroft, and Avi Levy. Finitely dependent cycle coloring. *Electronic Communications in Probability*, 23, 2018. doi:10.1214/18-ecp118.
- [44] Taisuke Izumi and François Le Gall. Quantum distributed algorithm for the all-pairs shortest path problem in the CONGEST-CLIQUE model. In Peter Robinson and Faith Ellen, editors, *Proceedings of the 2019 ACM Symposium on Principles of Distributed Computing, PODC 2019, Toronto, ON, Canada, July 29 - August 2, 2019*, pages 84–93. ACM, 2019. doi:10.1145/3293611.3331628.
- [45] Taisuke Izumi, François Le Gall, and Frédéric Magniez. Quantum distributed algorithm for triangle finding in the CONGEST model. In Christophe Paul and Markus Bläser, editors, *37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020, March 10-13, 2020, Montpellier, France*, volume 154 of *LIPICs*, pages 23:1–23:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPICS.STACS.2020.23.
- [46] Amos Korman, Jean-Sébastien Sereni, and Laurent Viennot. Toward more localized local algorithms: removing assumptions concerning global knowledge. *Distributed Computing*, 26(5-6):289–308, 2013. doi:10.1007/S00446-012-0174-8.
- [47] François Le Gall and Frédéric Magniez. Sublinear-time quantum computation of the diameter in CONGEST networks. In *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing*, pages 337–346. ACM, 2018. doi:10.1145/3212734.3212744.
- [48] François Le Gall, Harumichi Nishimura, and Ansis Rosmanis. Quantum advantage for the LOCAL model in distributed computing. In Rolf Niedermeier and Christophe Paul, editors, *36th International Symposium on Theoretical Aspects of Computer Science, STACS 2019, March 13-16, 2019, Berlin, Germany*, volume 126 of *LIPICs*, pages 49:1–49:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICS.STACS.2019.49.
- [49] Nathan Linial. Distributive graph algorithms-global solutions from local data. In *28th Annual Symposium on Foundations of Computer Science, Los Angeles, California, USA, 27-29 October 1987*, pages 331–335. IEEE Computer Society, 1987. doi:10.1109/SFCS.1987.20.
- [50] Nathan Linial. Locality in distributed graph algorithms. *SIAM Journal on Computing*, 21(1):193–201, 1992. doi:10.1137/0221015.
- [51] Frédéric Magniez and Ashwin Nayak. Quantum distributed complexity of set disjointness on a line. *ACM Transactions on Computation Theory*, 14(1):5:1–5:22, 2022. doi:10.1145/3512751.
- [52] Moni Naor and Larry J. Stockmeyer. What can be computed locally? *SIAM Journal on Computing*, 24(6):1259–1277, 1995. doi:10.1137/S0097539793254571.
- [53] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2012. doi:10.1017/cbo9780511976667.

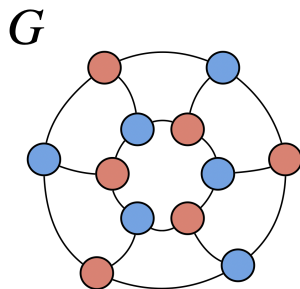


Figure 7: An example of a *lifebuoy-shaped* graph G . The lifebuoy-shaped graphs are defined by the set \mathcal{G} of graphs that are isomorphic to the above. Notice that their chromatic number is 2. In our case, the node labels required by the LOCAL model are unique and go from 1 to 12. Granting this extra power is not restrictive as we are proving a lower bound.

A Non-signaling and bounded-dependence models of distributed computing

This appendix introduces both the non-signaling model and the bounded-dependence model, which are the most powerful models that satisfy physical causality, and thus they generalize the quantum-LOCAL model. The distinction between the non-signaling model and the bounded-dependence depends on whether shared states are available (in the non-signaling model) or not (in the bounded-dependence model). We illustrate the difference between these models by analyzing the problem of $c = 2$ -coloring a lifebuoy-shaped graph (see Fig. 7 and Fig. 8b) with locality $T = 2$.

We first introduce a circuit formalism which allows us to clarify the early works of Arfaoui and Fraigniaud [3], and Gavioille et al. [33], by re-expressing the Randomized-LOCAL and Quantum-LOCAL models in this formalism (see Appendix A.1). Second, we introduce the concept of light-cones and the principle of non-signaling, explaining how (together with the symmetries of the graph) they define the non-signaling model (see Appendices A.1.3 and A.1.4). Third, we show in Appendix A.2 that lifebuoy-shaped graphs are not 2-colorable in the non-signaling model with locality $T = 2$ through a reduction from the 2-colorability of the cheating graph H shown in Fig. 8a. At last, we define the bounded-dependence model based on our circuit formalism and on the following three principles: device replication; non-signaling and independence [13, 24, 25, 38]; and invariance under symmetries of the graph (see Appendix A.3). We stress its connection to the concept of finitely dependent distributions [1, 40, 41, 43, 58].

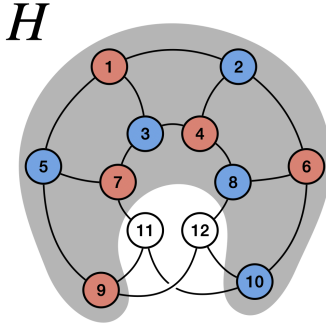
A.1 Randomized-LOCAL, quantum-LOCAL, and non-signaling models

In the next two sections, we re-introduce the Randomized-LOCAL model and the Quantum-LOCAL model for coloring lifebuoy-shaped graphs in a circuit formalism. This will later enable us to clarify the definitions of the non-signaling and bounded-dependence models.

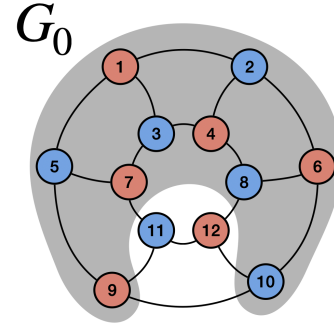
A.1.1 Randomized-LOCAL model

We consider twelve nodes with unique identifiers ranging from 1 to 12 and that are connected in a lifebuoy-shaped graph that is a priori unknown to the nodes³ (an example of such graph is the labeled graph $G_0 \in \mathcal{G}$ illustrated in Fig. 8b). The nodes try to color this graph in the randomized-LOCAL model within $T = 2$ steps of synchronous communication. The difficulty is that a given node ignore which $G \in \mathcal{G}$ connects

³A reader used to standard quantum nonlocality should see the graph — H or $G \in \mathcal{G}$ — as an input of the problem, split and distributed among the local parties.



(a) H has chromatic number 3.



(b) A distributed algorithm that finds a 2-coloring of the lifebuoy-shaped graphs $G \in \mathcal{G}$ would by definition 2-color the particular instance $G = G_0$.

Figure 8: The joint view of the couple of nodes $u, v = 1, 2$ after $T = 2$ rounds of communication is limited to the gray area. In this region the graphs G_0 and H are identical. The non-signaling principle implies that the outputs (c_1, c_2) must therefore be identically distributed in both G_0 and H .

the set of labelled nodes: it can only discover a local part of the structure of G by communicating to its neighbors. The most general ($T = 2$)-round strategy for the node 1 consists of the following procedure, which alternates between randomized processing steps and communication steps (the computational power and size of exchanged messages are unbounded):

Processing 0: Sample a random real number and store it locally.

Communication 1: Send all stored information, including the sampled random number, to all neighbors. Receive information from all neighbors and store it for subsequent rounds (in G_0 , the neighbors are 2, 3, 5).

Processing 1: Process all stored information (possibly in a randomized way) and store the result.⁴

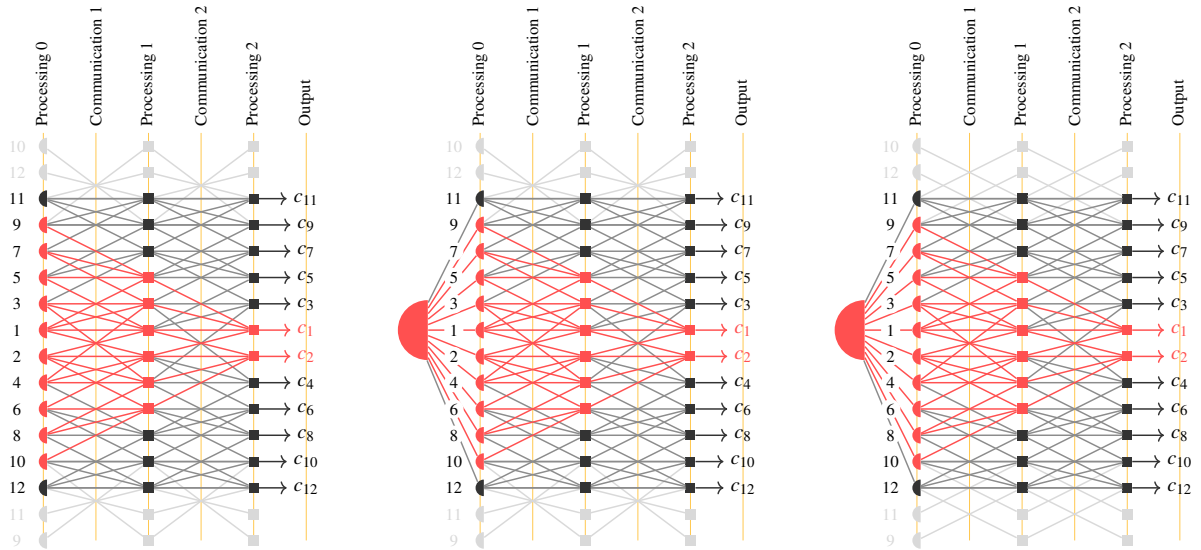
Communication 2: Send all stored information, including all received messages and the outputs of processing steps, to all neighbors. Receive information from neighbors and store it.

Processing 2: Process all stored information (possibly in a randomized way) to output a color.

Such T -round strategy on the graph G can be represented formally as a circuit $C_{G,T}$, such as in Fig. 9a where semicircles, line wires, and squares respectively represent the sampling of a random number, the transfer (or storage) of information, and the processing of information. Once the concrete operations performed by the nodes (i.e. randomness sampling and processing) are made explicit, it is possible to compute (using classical information theory) the exact output distribution of the strategy on graph G , that is, the probability distribution $\Pr[c_1, \dots, c_n \mid C_{G,T}]$ of observing that the set of nodes $\{1, \dots, n\}$ outputs the colors $\{c_1, \dots, c_n\}$ when connected as per one of the lifebuoy-shaped graphs $G \in \mathcal{G}$. Importantly, in our model, the operations performed by the nodes cannot depend on the connection graph G .

A generic classical strategy with shared randomness can be in the same way represented by the general circuit of Fig. 9b, by initializing the circuit with a source of randomness common to all nodes (the large semicircle).

⁴In classical information theory, it is known that intermediate processing gates can be taken as identity gates, that is, any strategy can be simulated by a two-layer circuit where the parties send their first random number to all parties up to a distance T , and then



(a) Circuit representation of a non-signaling strategy on graph G_0 of Fig. 8b, without a shared resource. Note the cyclicity of the circuit. Highlighted in red is the past-light-cone of the joint output (c_1, c_2) , that is the set of gates which connects to the output gates producing (c_1, c_2) .

(b) Circuit representation of a non-signaling strategy on graph G_0 of Fig. 8b, with an arbitrary shared resource. The joint output (c_1, c_2) remain, after $T = 2$ rounds of communication, independent of the graph structure around nodes 11 and 12, because the difference lies outside their joint past light-cones.

(c) Circuit representation of a non-signaling strategy on graph H of Fig. 8a, with an arbitrary shared resource. Note the difference between this circuit and the one of Fig. 9b (namely, the different connections between the top layer and bottom layer of the circuit) is only manifested outside the joint past light-cone of the outputs (c_1, c_2) .

Figure 9: The above circuits represent non-signaling strategies (it includes as special cases the classical strategies and quantum strategies) executed by the nodes 1 and 2 in various scenarios. The semicircles represent private (Fig. 9a) or shared (Fig. 9b and Fig. 9c) arbitrary-but-non-signaling resources; the wires depict communication (or storage), and the squares are local operations (using possibly private resources). The last layer of gates (or measurements) outputs the individual colors of the nodes (i.e. classical variables). Since the nodes start with no knowledge about the identity of their neighbors, the operations of the gates are a priori independent of the graph structure (as long as the nodes have the right degree). For the special cases of classical and quantum strategies, the output distribution can be computed directly from those circuits, which define it uniquely.

A.1.2 Quantum-LOCAL model

Any quantum strategy can also be represented by the circuit formalism of Fig. 9 by using resources and processing operations that are quantum rather than classical. Quantum information theory provides a concrete mathematical formalism (based on the tensor product of Hilbert spaces) with rules to represent the potential operations performed by each of the gates of the circuit. Semicircles, line wires, and squares now respectively represent the creation of a quantum state (a density matrix), the transfer (or storage) of quantum information, and the processing of quantum information with quantum channels (or quantum measurements in the last layer) represented by completely positive operators. Once the quantum states and channels are made explicit, the Born rule of quantum information theory [14, 53] allows computing $\Pr[c_1, \dots, c_n | C_{G,T}]$, the output distribution of the strategy on graph G .

However, analyzing the quantum-LOCAL model directly is complicated. Instead, we employ the fact that it is possible to bound the time complexity of the quantum-LOCAL model by analyzing models that do not depend on the mathematical formalism of quantum information theory (and which are, therefore, arguably simpler). In the present appendix, we utilize the fact that quantum-LOCAL is sandwiched between the randomized-LOCAL model (less powerful) and the non-signaling model (the most powerful model satisfying the information-theoretic principles of non-signaling, independence and device replication we introduce below). We next introduce the non-signaling model.

A.1.3 The non-signaling model (with unique identifiers)

The non-signaling model is the most powerful model of synchronous distributed computing that does not violate physical causality, when a pre-established shared resource exists.

The formulation of the non-signaling model is radically different from the randomized LOCAL and quantum-LOCAL models. The randomized LOCAL and quantum-LOCAL models are based on classical (randomized) and quantum information theory which both provide a mathematical formalism to describe information processing gates and to compute the probability distribution of outputs when these gates are placed in a circuit. To show that some probability distribution is feasible in these models, one needs to propose a valid strategy in these mathematical formalisms. The non-signaling model does not rely on any underlying mathematical formalism, but on some information-theoretical principles which should not be violated. More precisely, to show that some probability distribution is feasible in the non-signaling model, one needs to explain how to obtain it, but only to make sure that this probability distribution is compatible with the *non-signaling* principle in a way we now explain.

Before introducing this principle, we state the following definition of light-cones:

Definition A.1 (Past light-cone). Consider a circuit. Consider a subset of gates R , and another gate $s \notin R$ in that circuit. We say that s is in the past light-cone of R if starting from s , one can reach a gate in R by traveling down the circuit. For instance, in Fig. 9b, the past light-cone of the second processing gates of nodes $\{1, 2\}$ is composed of all gates in red.

Remark A.2. Light-cones allow to formalize the fact that, in a circuit, the precise time ordering in which the gates process information has no influence over the result, as long as physical causality (that is, information can be transferred only through communication) is preserved. For instance, in Fig. 9a, our drawing of the circuit represents the processing-1 gate of node 8 as being in the past of the processing-2 gate of node 2. However, by stretching the communication lines, another drawing in which this time ordering is exchanged exists, and this is possible as long as one node is not in the past light cone of another. The name “light-cone”

make a unique processing step after all the communication has taken place. In quantum and non-signaling theories, this is not the case anymore [27].

refers to relativity, where physical causality is bounded by the speed of light. Circuits provide an abstracted representation of physical causality that is not based on any notion of spacetime.

We now state the non-signaling principle⁵, which is essentially the only constraint in the non-signaling model.

Definition A.3 (Non-signaling principle). Consider a set of gates, two different circuits C, C' connecting them, and corresponding distributions $\Pr[c_1, \dots, c_n | C]$, $\Pr[c_1, \dots, c_n | C']$. Let U be a subset of measurement gates whose past light-cones in C and C' coincide (i.e. they are composed of the same gates which are connected in the same way). Then $\Pr[\{c_i\}_{i \in U} | C] = \Pr[\{c_i\}_{i \in U} | C']$.

Implicitly, a non-signaling theory assumes that, given a set of gates and a valid way to connect them in a circuit C , one obtains a corresponding distribution $\Pr[c_1, \dots, c_n | C]$ ⁶ (here the alphabet of the outputs c_i is not limited: e.g., in case of unexpected circuits, the measurement gates can produce failure outputs).

We are now ready to define the non-signaling model, which is associated to circuits with a pre-shared non-signaling resource such as in Fig. 9b:

Definition A.4 (Non-signaling model, with unique identifiers). The distribution

$$\Pr[c_1, \dots, c_n | C_{G,T}]$$

is feasible in the non-signaling model (with unique identifiers) with locality T on graph G if and only if, for all possible alternative connecting graphs H of the nodes, there exists a distribution $P = \Pr[c_1, \dots, c_n | C_{H,T}]$ such that the no-signaling principle is respected.⁷

Note that H might not be a lifebuoy-shaped graph: in the case a gate ‘discovers’ this unexpected situation, it has the possibility to produce a new output from a set of error outputs (i.e. it crashes and no further useful constraint can be derived).

A.1.4 The non-signaling model without unique identifiers

We have up until now considered $N = 12$ nodes with unique identifiers ranging from 1 to $N = 12$. When the nodes do not start with unique identifiers, the situation is slightly more complicated. In that case, the resulting distribution should be invariant under subgraph isomorphism, as all nodes are running identical programs and the circuit is thus completely symmetric under permutations of nodes⁸. For instance, if the nodes were to color the graph of Fig. 8b without being a priori assigned unique identifiers, the resulting circuit $C_{G,T}$ would have some symmetries (as all processing gates would be the same in any layer of the respective steps 0, 1 and 2), implying that the distribution should be invariant by several non-trivial graph isomorphisms cyclically permuting the nodes, or inverting the inner and outer cycle of 6 nodes.

⁵Chiribella and Spekkens [23] call this principle *no-signaling from the future*.

⁶Note that the gates have “types” and can only be connected to other gates of matching types: For instance, in our case, the inputs of gates from the second processing step must correspond to the outputs of gates from the first processing step. The theory of circuits can be formalized using category theory — see, e.g. [30].

⁷Note that considering scenarios with more than one copies of each node does not allow us to derive additional constraints, since the pre-shared non-signaling resource cannot by definition be cloned and extended to more than one copy of each of its original recipient.

⁸Note that while a shared classical resource is inherently symmetric, because it can be without a loss of generality taken to be distributed identically to each party, a non-signaling resource is not a priori symmetric under permutation of the parties that it connects. The absence of identifiers thus forces the non-signaling model to pre-share only a subset of all possible non-signaling resources, namely the subset that respects such symmetry.

Definition A.5 (Non-signaling model, without unique identifiers). The distribution

$$\Pr[c_1, \dots, c_n \mid C_{G,T}]$$

is feasible in the non-signaling model (without unique identifiers) in T communication steps on graph G if and only if the following two conditions are respected: for all possible alternative connecting graphs H of the nodes, there exists a distribution $P = \Pr[c_1, \dots, c_n \mid C_{H,T}]$ such that the no-signaling principle is respected; and the distributions in G and H are invariant under subgraph isomorphism.

This definition can also be generalized to the case where several nodes with the same identifiers are present in G , or where the number of identifiers ranges from 1 to $M \neq N$.

A.2 Graph \mathcal{G} is not 2-colorable in $T = 2$ rounds

We now show by contradiction that the nodes $1, \dots, 12$ cannot 2-color the set \mathcal{G} of lifebuoy-shaped graphs with a non-signaling strategy (including pre-shared non-signaling resources) in $T = 2$ rounds of communication. More precisely, we prove that if there existed such a non-signaling algorithm able to color any $G \in \mathcal{G}$, then the same algorithm would also color the cheating graph H with 2 colors, which is impossible.

Proof. Assume by contradiction that there exists such a non-signaling algorithm, that is, there exist gates as in Fig. 9b such that the resulting probability distribution $\Pr[c_1, \dots, c_{12} \mid C_{G,T}]$ is a proper coloring for all lifebuoy-shaped graphs $G \in \mathcal{G}$. Such proper coloring means that for all $G \in \mathcal{G}$, and for all nodes u, v that are neighbors in G , it holds that $\Pr[c_u \neq c_v \mid C_{G,T}] = 1$.

We consider the same nodes performing the same gates, but we change the edges so that the connected nodes now form the graph H in Fig. 8b. Let $\Pr[c_1, \dots, c_{12} \mid C_{H,T}]$ be the distribution of output colors in that new configuration⁹. Consider two nodes connected by an edge in H : by symmetry of H , we can without a loss of generality assume that these are nodes 1 and 2. We introduce the lifebuoy-shaped graph G_0 of Fig. 8b ($G_0 \in \mathcal{G}$ depends on our choice of nodes, here 1, 2, in H). Then, we observe that the common past light-cones of the two nodes is the same when connected through H or through G_0 (compare the circuit of Fig. 9b with the circuit of Fig. 9c). It hence holds that $\Pr[c_1 \neq c_2 \mid C_{H,T}] = 1 = \Pr[c_1 \neq c_2 \mid C_{G_0,T}]$ due to the non-signaling principle. We conclude that the non-signaling distributed algorithm outputs a 2-coloring for H , which is impossible. \square

A.3 Bounded-dependence model

The bounded-dependence model is similar to the non-signaling model, but the former does not allow pre-shared non-signaling resources between the nodes. The following two new principles are needed to define feasible distributions in the bounded-dependence model.

Definition A.6 (Independence principle). Consider a set of gates connected in a circuit C , and the corresponding distribution $\Pr[c_1, \dots, c_n \mid C]$. Let U, V be two subsets of measurement gates producing the outputs $\{c_u\}_{u \in U}$ and $\{c_v\}_{v \in V}$, respectively. If the past light-cones of U and V do not intersect, then $\Pr[\{c_w\}_{w \in U \cup V} \mid C] = \Pr[\{c_u\}_{u \in U} \mid C] \cdot \Pr[\{c_v\}_{v \in V} \mid C]$, that is their output distributions are independent.

Definition A.7 (Device-replication principle). Identical and independent copies of non-signaling gates and non-signaling resources can be prepared¹⁰.

⁹While we are deceiving the nodes by promising a lifebuoy-shaped connecting graph but imposing instead H , the nodes cannot locally detect the fraud in $T = 2$ communication steps or less. (More formally, the past light-cone in H of any node is then compatible with a lifebuoy-shaped graph — an individual node cannot detect the difference and must therefore, according to the non-signaling principle, output a color as if it were in a lifebuoy-shaped graph.)

¹⁰Note that this principle does not imply that one can duplicate *unknown* non-signaling resources: device-replication compatible with the quantum no-cloning theorem.

Then, we obtain the following definition:

Definition A.8 (Bounded-dependence model, with unique identifiers). The distribution

$$\Pr[c_1, \dots, c_n \mid C_{G,T}]$$

with unique identifiers has *bounded dependence* with locality T on graph G (without pre-shared non-signaling resources) if and only if for all possible alternative connecting graph H of the nodes and their replicates, there exists a distribution $\Pr[c_1, \dots, c_m \mid C_{H,T}]$ such as the non-signaling and independence principles are respected, and such that the distribution is invariant under subgraph isomorphisms.

Note a subtlety related to subgraph isomorphisms in the bounded-dependence model. There is the variant where the nodes have identifiers in G , and the one where they do not. When the nodes do not have any identifiers, the class of subgraph isomorphisms of all alternative graphs H created out of the nodes in G and their replicates is obviously larger than with identifiers, because the subgraph isomorphisms must respect the identifiers. However, even if the nodes of G do have distinct identifiers, as H is created out of possibly many copies of the original nodes of G , H might contain several nodes with the same identifiers. Hence, the group of subgraph isomorphism of H might be nontrivial even if all nodes in G have distinct identifiers. For instance, in lifebuoy-shaped graphs, one could consider the case represented in Fig. 10, which starts from the graph G_0 in Fig. 8b, duplicates all nodes, and constructs a new graph H of 24 nodes with identifiers ranging from 1 to 12 with one non-trivial graph isomorphism cyclically permuting the nodes.

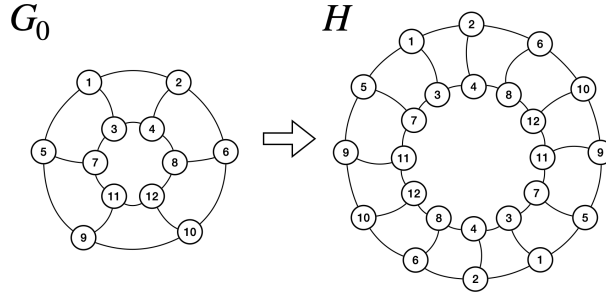


Figure 10: In the bounded-dependence model, one can find non-trivial subgraph isomorphisms even when the nodes are provided with unique identifiers.

A.3.1 Relation with finitely dependent distributions

Our bounded-dependence model is directly connected to the concept of finitely dependent distributions introduced in mathematics [1, 40–43, 58]. In this framework, in a graph H , the color c_i produced by each node i is seen as the result of a random process C_i . The set of all random processes $\{C_i\}_i$ is said to be k -dependent in graph H if, for any two subsets U, V of the nodes of H which are at least at distance $k + 1$, the two sets of processes $\{C_u\}_{u \in U}$ and $C_{v \in V}$ are independent. In our notation, assuming even k and taking $T = k/2$, this is equivalent to asking that $\Pr[\{c_w\}_{w \in U \cup V} \mid C_{H,T}] = \Pr[\{c_u\}_{u \in U} \mid C] \cdot \Pr[\{c_v\}_{v \in V} \mid C_{H,T}]$.

Hence, the question of the existence of distributions with bounded dependence that solve some problem can directly be formulated in terms of the existence of finitely dependent distributions over a family of graphs that satisfy additional constraints of compatibility (to satisfy the non-signaling principle) and symmetries (to cope with subgraph isomorphisms).