Distributed Computing Column 55

WRAWN 2013 Review, WTTM 2013 Review, and Lower Bounds for Distributed Quantum Computing

Jennifer L. Welch
Department of Computer Science and Engineering
Texas A&M University, College Station, TX 77843-3112, USA
welch@cse.tamu.edu

In the first part of this issue's column, Magnús Halldórsson and Calvin Newport take a fresh perspective on doing a workshop review. They highlighted five big-picture ideas that came out of the presentations and discussion at 2013 Workshop on Realistic Models of Wireless Networks (WRAWN) which they organized and was co-located with PODC. The focus of the workshop was whether wireless theory is sufficiently relevant to wireless practice, and if not, what can be done to improve matters.

The second part of the column is a review of the 2013 Workshop on the Theory of Transactional Memory, which was co-located with DISC. The authors, Claire Capdevielle and Sandeep Hans, have done a great job of summarizing the talks as well as putting the topics of the talks into the wider context of prior work.

The column concludes with a research paper by Heger Arfaoui and Pierre Fraigniaud on lower bounds for *quantum* distributed algorithms. The beauty of their approach is that the bounds can be shown without having to manipulate concepts from quantum mechanics at all! A key idea, as in the DISC 2009 paper by Gavoille, Kosowski, and Markiewicz, is to identify a class of distributions that subsumes those of quantum computing but that is easier to work with. This is then exploited to show that for most two-player games in which the players cannot communicate, quantum correlations do not allow a higher probability of success than the classical model.

Many thanks to Magnús, Calvin, Claire, Sandeep, Heger and Pierre for their contributions!

Call for contributions: I welcome suggestions for material to include in this column, including news, reviews, open problems, tutorials and surveys, either exposing the community to new and interesting topics, or providing new insight on well-studied topics by organizing them in new ways.

Making Wireless Algorithm Theory More Useful: Five Ideas from the 2013 Workshop on Realistic Models for Algorithms in Wireless Networks

Magnús Halldórsson
Reykjavik University
Reykjavik, Iceland
mmh@ru.is

Calvin Newport
Georgetown University
Washington D.C., USA
cnewport@cs.georgetown.edu

1 Overview

In summer 2013, a group of invited speakers and interested participants gathered in Montreal, Canada, for the fourth annual Workshop on Realistic Models for Algorithms in Wireless Networks (WRAWN), co-organized by the authors of this article, and co-located with the ACM Principles of Distributed Computing (PODC) conference. The motivating question for this year's workshop was blunt:

Is the study of wireless algorithms useful enough?

To help explore answers, the workshop featured talks that highlighted different perspectives on this question.[1] Fabian Kuhn and Christian Scheideler discussed wireless models that include non-determinism, boosting the robustness of algorithmic results, while Seth Gilbert motivated the need to study multi-channel networks. Majid Khabbazian and Thomas Janson discussed the theory and practice of interference cancellation strategies such as analog network coding and MIMO, while Magnús Halldórsson and Martin Hoefer described efforts to identify the fundamental wireless properties needed to prove useful results. Finally, Nancy Lynch talked about new settings, such as biological systems, where wireless distributed algorithm theory can play a useful role, and Calvin Newport explored the current gap between wireless theory and practice.

These talks fueled free-form discussions on the current state of wireless algorithm theory. We debated which research directions were useful and which were dead ends, and we attempted to better understand the connection (or lack thereof) between this work and practice. Though the

[1] Talks slides are available at the WRAWN 2013 web site: http://wrawn.ru.is/program13.htm.

workshop generated many useful insights, in this article we summarize *the five most interesting ideas from these discussions*.[2] These ideas do not present a comprehensive road map for the future of wireless algorithm theory, but they will hopefully stimulate new ways to think about the role of theoreticians in the development of this increasingly important technology.

Idea #1: Recognize the Growing Necessity for Wireless Algorithm Theory. There currently exists a significant gap between wireless algorithm theory and practice. Engineers have been building these networks since the 1970s, and theoreticians have been studying algorithms in this setting since the 1980s, but few results during this period have moved from theory to practice. The *meta-idea* that drives those that follow is that this state of affairs is changing. There are three explanations for this trend. The first explanation is the long-awaited realization of ubiquitous computing. For years, technologists have preached a future in which networked computational devices infuse our environment, working together to improve our lives. Most of these devices would be networked wirelessly, and due to power constraints and sparsity of spectrum, many will likely rely on local low-power links to form ad hoc networks. This is a setting where distributed algorithm theory can play a large role. In recent years, this vision is suddenly becoming viable, thanks in large part to the rise of the smartphone industry, which has advanced wireless and battery technology while simultaneously driving down the prices of these components. A second trend is the recent move to expand existing wireless standards such as 802.11 and Bluetooth to better support the development of applications that use local wireless communication. Many of the applications running on these local wireless networks will benefit from distributed algorithms (imagine, for the example, the need for an efficient replicated state machine algorithm so a group of nearby laptops can run a multiplayer game uninterrupted, even as users come and go). The final trend is that the rising performance and falling cost of wireless connectivity will lead to an increasing number of traditionally wired distributed systems to replace wires with wireless links. (Such systems can be cheaper to deploy and easier to manage.) The distributed algorithms used by these systems will have to be redesigned and analyzed for the wireless communication medium.

Idea #2: Generalize as Much as Possible (But No More). Much of the existing work on algorithms for wireless networks probes the *basic science* of the setting; e.g., identifying fundamental limits for contention management among unknown contenders, or bounding the capacity of fading channels. In pursuing these basic science questions it is important to keep generalizing the key properties of one's model until arriving at the threshold where useful results are no longer possible. When studying scheduling problems in signal strength models, for example, many powerful results are possible when the nodes are embedded in a general metric space, but when moved to arbitrary metrics, these same problems become intractable. This threshold helps identify exactly what properties of fading enable efficient spatial refuse of the medium. Answers to such questions are useful as they provides a deeper insight into wireless networks—insight that will be increasingly important as our dependence on this technology grows.

Idea #3: Embrace Non-Determinism. Many of the existing models used to study wireless algorithms depend on fixed deterministic rules for describing the underlying network behavior; i.e., a given set of senders uniquely determines the receive behavior. This property is true of

[2]We emphasize that the selection and wording of these ideas is due solely to this article's authors, and do not necessarily represent the views of the other workshop participants.

both the standard graph and signal strength models. Real wireless networks, by contrast, are less predictable, due to the fact that the wireless medium is open and shared. Simplified deterministic rules makes sense when studying basic science problems of the type discussed in Idea #1. They might cause a problem, however, when these models are used to design and analyze algorithm meant for deployment. A property carefully proved in a model with fixed behavior, for example, might fail when the algorithm is deployed in a real network where signals do no always fade cleanly, or concurrent broadcasts do not always create collisions. The study of algorithms has a well-known answer to this problem: introduce non-determinism into your models. In a graph setting, for example, you might allow the topology to change, while in a signal strength setting, you might allow the distance metric to vary. Finding good definitions for these models, understanding their limits, and identifying algorithmic strategies that tame their unpredictability, are research directions that are beginning to gain steam within our community. This emphasis should be encouraged.

Idea #4: Explore the Stack. The vast majority of algorithms studied for wireless networks execute at the link layer of the network stack. This means that the algorithms are responsible for standard link layer activities such as managing contention on the shared channel. This perspective is useful for investigating basic science questions (see Idea #2), and for aiding the development of new link layer strategies, as might be needed by the growing diversity of wireless devices (see Idea #1). The rise of shared spectrum networking, for example, requires a whole new multi-channel perspective on contention management. Theoreticians in our community, however, should spend more time studying the *other* layers of the network stack—each of which offers opportunities for theory to solve practical problems. At the physical layer, for example, the introduction of interference cancellation strategies such as MIMO induce many theory-style questions that must be answered to broaden these technologies' usefulness. Jumping up to the network layer, the algorithmic designer now faces a model that abstracts the guarantees of existing link layers—freeing the designer to focus on proving higher-level properties. Algorithms designed and analyzed at this level of abstraction are especially applicable to practice as they can be deployed on existing devices running existing link layer standards. Even the application layer offers new challenges for theoreticians. Due to the mobility of the wireless setting, for example, applications in such networks must maintain key properties in the face of unusually dynamic membership, another setting where distributed algorithms are relevant.

Idea #5: Tell Your Story. The theory community studying wireless algorithms has generated, over its thirty-year history, a powerful collection of fundamental limits and algorithmic strategies. Much of this work has direct relevance to real network settings. We have developed, for example, simple randomized contention management strategies (based on the cyclical testing of different contention levels), that are fully distributed and can eliminate the hidden terminal problem that plagues the CSMA-style strategies used in most deployed wireless devices. Similarly, we have a large corpus of results on the relationship between power control and the capacity of the wireless medium—work that indicates that standard approaches to power (such as using uniform or linear power assignments) can leave a massive amount of communication capacity untapped. Our community needs to tell these stories. Here are three things that might aid this effort: (1) give theory papers practitioner-parsable titles; (2) make *impact on practice* a standard part of the front matter in our papers; (3) give talks to networking audiences (then listen to their feedback).

WTTM 2013, The Fifth Workshop on the Theory of Transactional Memory

Claire Capdevielle
University of Bordeaux
France
ccapdevi@labri.fr

Sandeep Hans
Technion
Israel
sandeep@cs.technion.ac.il

Abstract

In conjunction with DISC 2013, the TransForm project (Marie Curie Initial Training Network) and EuroTM (COST Action IC1001) supported the 5th edition of the Workshop on the Theory of Transactional Memory (WTTM 2013). The objective of WTTM was to discuss new theoretical challenges and recent achievements in the area of transactional computing with emphasis on transactional memory. The workshop took place on October 14, 2013, in Jerusalem, Israel. This report is intended to give highlights of the problems discussed during the workshop.

1 Disjoint-access-parallelism

Disjoint-access-parallelism (dap) seems to be a fundamental property to ensure scalability of Transactional Memory (TM) implementations, since it formalizes the idea that transactions on disjoint data should not interfere. A TM implementation is an algorithm that implements transactions by applying primitive operations on base objects. Several variants of disjoint-access-parallelism have been proposed since its introduction by Israeli and Rappaport [16]. Some of them are formalized through the notion of conflict graph: a conflict graph is a graph whose nodes represent transactions and an edge exists between two nodes if the two transactions access the same data item. Strict disjoint-access-parallelism requires that two processes p and q performing two transactions T_1 and T_2 access the same base object only if there is an edge between T_1 and T_2 in the corresponding conflict graph. On the other hand, according to weak disjoint-access-parallelism p and q can access a common base object only if there is a path between T_1 and T_2 in the graph.

In [10] Guerraoui and Kapalka demonstrated that no obstruction-free, serializable TM can be strict-disjoint-access-parallel. In her first talk, **Panagiota Fatourou** investigated if an obstruction-free TM can ensure strict disjoint-access-parallelism and snapshot isolation [2]: a consistency criterion weaker than serializability. The response is still negative. Then, she presented an obstruction-free transactional memory, called SI-DSTM, which implements snapshot isolation and ensures strict

disjoint-access-parallelism for read-only transactions and weak disjoint-access-parallelism for writing transactions.

The second talk by **Panagiota Fatourou** was about circumventing another impossibility on disjoint-access-parallelism: no universal construction can ensure wait-freedom and disjoint-access-parallelism [7]. The universal construction is a paradigm close to the TM: it provides a general mechanism for obtaining a concurrent implementation of any object from its sequential code. The above impossibility is valid for all the variants of dap proposed in the literature. Also it holds for transactional memories where transactions can abort only a finite number of times.

Panagiota Fatourou showed that the impossibility proved in [7] can be overcome if we consider a variant of dap, called *timestamp-ignoring disjoint-access-parallelism*, which is similar to weak dap but allows transactions to access a wait-free timestamp object, even though they access disjoint sets of data items.

2 On Concurrent Objects Implementation

The keynote by **Michel Raynal** was about concurrent objects implementations. He reviewed the major existing approaches: lock-based implementations where the accesses to the objects are synchronized using locks, and lock-free implementations which avoid locks or similar mechanisms. Coarse-grain lock-based implementations are easy to program and verify, but cannot cope with failures and asynchrony. On the other hand, implementations that do not use locks cope with failures and asynchrony but are usually difficult to conceive and may impose too much overhead. He suggested hybrid implementations which mix the above approaches.

The implementation of a concurrent object is static hybrid if some of its operations are implemented with locks while others are not. As an example, from [14] he presented a hybrid static implementation of a concurrent set object S manipulated through *add* (add an element to S), *remove* (remove an element from S) and *contain* operations (search for an element in S). The algorithms that implement the *add* and *remove* use locks while the one implementing the *contain* operation does not. This choice for synchronization is interesting when it is assumed that the contain operation is invoked much more than the others.

He also presented dynamic hybrid implementations where locks are used according to the context. The main idea is not to use locks in "favorable circumstances", where favorable circumstances usually meant lack of concurrency. But according to Raynal's talk this notion can be generalized. For example, a favorable circumstance for a consensus object is when all the processes propose the same value [21].

At the end of his talk, he also discussed the notion of abortable object. An object is abortable if any invocation of an operation on this object returns after a bounded number of steps and is allowed to return abort in the presence of concurrency. An abort indicates that the operation did not take effect. Abortable objects can be implemented with hybrid solutions.

3 Contention Management: From Multiprocessor to Distributed TMs

Danny Hendler's keynote talk was about contention management in software transactional memory (STM). Contention management is present in transactional memories to decide, in case of

contention, which transaction must abort and when to re-execute it. Several policies have been proposed in the context of multiprocessor architectures [20], but none of these has been proved to be efficient under high contention. Therefore, another mechanism, the TM scheduler, has been proposed to prevent conflicts between transactions.

After presenting a few schedulers conceived for multiprocessor TM, Danny Hendler discussed the challenges of implementing transactional schedulers in a distributed setting. Distributed transactional memory (DTM) extends the TM abstraction to distributed applications. DTM has a larger design space than TM: transactional data can be stored at fixed nodes and transactions move from node to node to be executed (control flow model); or data is moved locally to nodes that have to execute transactions (data flow model). Also, replication is often used to support concurrent execution of transactions at different nodes. Thus contention management in DTM consists of a local contention manager that decides about transactions invoked on the same node, and a remote contention manager that resolves conflicts between transactions executed concurrently at different nodes. For the local contention management, current DTMs use the policies proposed in the context of TM without taking into account the policy of the remote contention management. The important point is to improve the synergy between them.

Shlomi Dolev presented SemanticTM, an opaque TM algorithm that avoids aborting transactions. To this aim each transactional variable is associated to a list. A thread (the scheduler) places the instructions of each transaction with their dependencies in the appropriate list before the instructions of any subsequent transaction. Other threads execute the instructions from the lists in order. The algorithm guarantees fine-grained parallelism, i.e., the parallelism is at the transactional instruction level and not at the transaction level.

Costas Busch's talk was about three analogies between schedulers in multiprocessor systems and network problems. This was joint work with Gokarna Sharma.

The first parallel was between packet scheduling techniques and transaction scheduling in multicores. More precisely, with the inspiration of [18], he made an analogy between packets and threads, path length and sequence of thread's transaction, and network congestion and conflicts of thread's transactions. He proposed a method to schedule M threads with a sequence of N transactions, with a makespan (time to complete all transactions) in $O(C + N \log(MN))$, assuming that each transaction has conflicts with at most C other transactions.

The second parallel was between mobile object tracking in sensor networks and scheduling of transactions in networked systems. With the inspiration of [1], he presented Spiral, a new directory protocol for the data flow model. A directory protocol supports the move of transactional objects to nodes that want to access them. Spiral allows a stretch of $O(\log(n)^2 \log(D))$ where n is the number of nodes in the network and D the diameter of the network.

Finally the third parallel was between oblivious routing in networks and load balance transactional scheduling in NUMA.

Jean-Philippe Martin presented his joint work with Christopher J. Rossbach on a model that predicts the performance for distributed software transactional memory (DSTM) executing a given workload. The model is validated in comparison with real executions from TM benchmarks, then it is applied to popular TM benchmarks on DSTM. He showed that the current TM benchmarks are not appropriate workloads for DSTM that rely on speculation and thus have an optimistic approach to manage concurrency. The reason for this is that the transactions are too short in relation to network latencies. Thus, new benchmarks are needed to test the DSTM using optimistic concurrency.

4 TM Semantics

In her keynote, **Hagit Attiya** suggested a new approach for evaluating and comparing TM consistency conditions. It can also reduce the effort of proving that a TM implements its programming language interface correctly, by only requiring its developer to show that it satisfies the corresponding consistency condition. The motivation of this talk was to close the gap between given consistency conditions and formalizing the intuitive semantics of atomic blocks from a programmer's perspective. The talk showed that, for a particular programming language and notions of observable behavior, a variant of the well-known consistency condition of opacity [11] is sufficient for observational refinement [12, 13], and its restriction to complete histories is furthermore necessary.

Jens Palsberg presented intuitive invariants for the correctness of TM algorithms. A history is markable if there is a specific ordering relation called marking such that three invariants are satisfied. These invariants are not only required but also sufficient for opacity. He proved the equivalence of markability and opacity. Roughly speaking, the first invariant, called write-observation, requires that each read operation returns the most current value; the second invariant, called read-preservation, requires that the location which is read is not overwritten in a certain interval; and the third invariant is the well-known real-time-preservation property.

Most transactional memory specifications do not consider local and non-transactional operations. However, systems provide a variety of synchronization mechanisms, and TM must be able to interact with them. Therefore, a specification for TM must specify the interaction with non-transactional operations also. Prior work on verifying TM with non-transactional operations required "Strong Atomicity", which says that non-transactional operations are equivalent to "mini-transactions" that cannot abort. TMS1 is a specification given by Lesani et al. They extended TMS1 [6] to get NTMS1 to allow non-transactional operations. In his talk, **Victor Luchangco** showed that a data-race-free program cannot distinguish a transactional memory implementation satisfying NTMS1 from one that provides strong atomicity. Thus, NTMS1 guarantees strong atomicity for data-race-free clients. However, NTMS1 does not guarantee privatization-safety.

Yujie Liu presented the complexities that arise by reordering instructions within transactions by the programmer or compiler. A programmer or a compiler can reorder the instructions within a transaction to improve instruction scheduling but it can give semantic problems. This can lead to frequent aborts, and can affect the progress. He showed a mechanism for delaying the conflicting transactional operations until commit time, where they cannot be aborted. It is assumed that the programmer will annotate the operators he wishes to reorder. Liu gave an example of how reordering violates publication safety. He also showed between such a mechanism and language-level semantics, mainly two restrictive levels of transactional semantics proposed by Menon et al. [19] - *Asymmetric Lock Atomicity* (ALA) and *Encounter-time Lock Atomicity* (ELA). He showed that their solution for ELA semantics ensures safety of privatization but requires versioned writes.

5 Hybrid TM

Alex Matveev presented a reduced hardware version of the NORec Hybrid TM [3] algorithm. It provides opacity with low hardware abort rates, and the mostly software slow-path is obstruction-free. It uses the short hardware transactions only to write values during the software commit. It overcomes the drawbacks of the earlier Hybrid NORec proposals by reading the shared global clock of the NORec STM [4] only at the end of the hardware transaction, thus providing opacity with

low hardware abort rates.

As programmers begin to write programs with transactions, particularly large transactions, scaling problems are inevitable, and we can expect growing demand for programming techniques that minimize transaction conflicts and hardware overflow. **Lingxiang Xiang** introduced one such technique. Unlike early release [15], elastic transactions [8] and composing relaxed transactions [9], partitioned transactions are compatible with existing HTM. They induce a model in which the programmer highlights the data that a transaction needs, rather than the data that it does not.

6 Multiversion Transactional Memories

The Time-Warping Multi-version (TWM) algorithm [5] was proposed to minimize spurious aborts without hampering practical performance. The extent to which a TM allows such spurious aborts is captured by the theory of Input Acceptance. **Nuno Diegues** compared TWM with other existing TMs using input acceptance. The key idea of input acceptance is to identify a sequence of input events (an input pattern) that, when fed to a TM, leads to aborting at least one transaction.

A commonly accepted correctness criterion for STM systems is opacity [11], proposed by Guerraoui and Kapalka. Opacity requires all the transactions (including aborted) to appear to execute sequentially in an order that agrees with the order of non-overlapping transactions. The main motivation of STM implementations is to increase concurrency to get better performance. But, with the increase in concurrency, more transactions may conflict and abort, especially in the presence of long transactions, which can deteriorate the performance drastically. It was observed that by storing multiple versions of each object, multi-version STMs can ensure that more read operations are successful, giving more concurrency and hence better performance. The correctness criterion in databases for multi-version is multi-version view-serializability (MVSR) [22], but checking for membership of MVSR has been proved to be NP-Complete. However, a sub-class of MVSR, conflict-serializability (CSR) has been identified, whose membership is easy to verify. Similarly, using the notion of conflicts, a sub-class of opacity, conflict-opacity (co-opacity) [17] is designed whose membership can be verified easily. The main drawback of co-opacity is that it does not admit histories that use multiple versions. **Sathya Peri** presented a new notion of conflict called multi-version conflict. Using this conflict notion, a new subclass of opacity, mvc-opacity, can be identified, whose membership is easier to verify. He further showed that co-opacity is a proper subset of this class.

The abstracts and slides of the talks can be downloaded at the following link:
http://www.gsd.inesc-id.pt/∼mcouceiro/wttm2013/html/index.html.

References

[1] Baruch Awerbuch and David Peleg. Concurrent online tracking of mobile users. *SIGCOMM Comput. Commun. Rev.*, 21(4):221–233, August 1991.

[2] Hal Berenson, Phil Bernstein, Jim Gray, Jim Melton, Elizabeth O'Neil, and Patrick O'Neil. A critique of ANSI SQL isolation levels. *SIGMOD Rec.*, 24(2):1–10, May 1995.

[3] Luke Dalessandro, Franois Carouge, Sean White, Yossi Lev, Mark Moir, Michael L. Scott, and Michael F. Spear. Hybrid norec: a case study in the effectiveness of best effort hardware transactional memory. In *ASPLOS*, pages 39–52. ACM, 2011.

[4] Luke Dalessandro, Michael F. Spear, and Michael L. Scott. Norec: Streamlining STM by abolishing ownership records. *SIGPLAN Not.*, 45(5):67–78, January 2010.

[5] N. Diegues and P. Romano. Brief announcement: Enhancing permissiveness in transactional memory via time-warping. In *DISC*, 2013.

[6] Simon Doherty, Lindsay Groves, Victor Luchangco, and Mark Moir. Towards formally specifying and verifying transactional memory. *Formal Aspects of Computing*, pages 1–31, March 2012.

[7] Faith Ellen, Panagiota Fatourou, Eleftherios Kosmas, Alessia Milani, and Corentin Travers. Universal constructions that ensure disjoint-access parallelism and wait-freedom. In *Proceedings of the 2012 ACM Symposium on Principles of Distributed Computing*, PODC '12, pages 115–124, New York, NY, USA, 2012. ACM.

[8] Pascal Felber, Vincent Gramoli, and Rachid Guerraoui. Elastic transactions. In *DISC*, pages 93–107, 2009.

[9] Vincent Gramoli, Rachid Guerraoui, and Mihai Letia. Composing relaxed transactions. In *IPDPS*, pages 1171–1182, 2013.

[10] Rachid Guerraoui and Michal Kapalka. On obstruction-free transactions. In *Proceedings of the Twentieth Annual Symposium on Parallelism in Algorithms and Architectures*, SPAA '08, pages 304–313, New York, NY, USA, 2008. ACM.

[11] Rachid Guerraoui and Michal Kapalka. On the correctness of transactional memory. In *PPOPP*, pages 175–184, 2008.

[12] Jifeng He, C. Hoare, and J. Sanders. Data refinement refined. In *ESOP*, pages 187–196, 1986.

[13] Jifeng He, C. Hoare, and J. Sanders. Prespecification in data refinement. *Information Processing Letters*, 25(2):71 – 76, 1987.

[14] Steve Heller, Maurice Herlihy, Victor Luchangco, Mark Moir, William N. Scherer III, and Nir Shavit. A lazy concurrent list-based set algorithm. In James H. Anderson, Giuseppe Prencipe, and Roger Wattenhofer, editors, *Principles of Distributed Systems*, volume 3974 of *Lecture Notes in Computer Science*, pages 3–16. Springer Berlin Heidelberg, 2006.

[15] Maurice Herlihy, Victor Luchangco, Mark Moir, and William N. Scherer III. Software transactional memory for dynamic-sized data structures. In *PODC*, pages 92–101, 2003.

[16] Amos Israeli and Lihu Rappoport. Disjoint-access-parallel implementations of strong shared memory primitives. In *Proceedings of the Thirteenth Annual ACM Symposium on Principles of Distributed Computing*, PODC '94, pages 151–160, New York, NY, USA, 1994. ACM.

[17] Petr Kuznetsov and Sathya Peri. On non-interference of transactions. *CoRR*, abs/1211.6315, 2012.

[18] F.T. Leighton, Bruce M. Maggs, and Satish B. Rao. Packet routing and job-shop scheduling in o(congestion+dilation) steps. *Combinatorica*, 14(2):167–186, 1994.

[19] Vijay Menon, Steven Balensiefer, Tatiana Shpeisman, Ali-Reza Adl-Tabatabai, Richard L. Hudson, Bratin Saha, and Adam Welc. Practical weak-atomicity semantics for Java STM. In *Proceedings of the Twentieth Annual Symposium on Parallelism in Algorithms and Architectures*, SPAA '08, pages 314–325, New York, NY, USA, 2008. ACM.

[20] William N. Scherer, III and Michael L. Scott. Advanced contention management for dynamic software transactional memory. In *Proceedings of the Twenty-fourth Annual ACM Symposium on Principles of Distributed Computing*, PODC '05, pages 240–248, New York, NY, USA, 2005. ACM.

[21] Gadi Taubenfeld. Contention-sensitive data structures and algorithms. In *Proceedings of the 23rd International Conference on Distributed Computing*, DISC'09, pages 157–171, Berlin, Heidelberg, 2009. Springer-Verlag.

[22] Gerhard Weikum and Gottfried Vossen. *Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control and Recovery.* Morgan Kaufmann Publishers Inc., 2001.

What Can Be Computed without Communications? [1]

Heger Arfaoui
CNRS and University Paris Diderot
heger.arfaoui@liafa.univ-paris-diderot.fr

Pierre Fraigniaud
CNRS and University Paris Diderot
Pierre.Fraigniaud@liafa.univ-paris-diderot.fr

Abstract

The main objective of this paper is to provide illustrative examples of distributed computing problems for which it is possible to design tight lower bounds for *quantum* algorithms without having to manipulate concepts from quantum mechanics, at all. As a case study, we address the following class of 2-player problems. Alice (resp., Bob) receives a boolean x (resp., y) as input, and must return a boolean a (resp., b) as output. A *game* between Alice and Bob is defined by a pair (δ, f) of boolean functions. The objective of Alice and Bob playing game (δ, f) is, for every pair (x, y) of inputs, to output values a and b, respectively, satisfying $\delta(a, b) = f(x, y)$, in *absence of any communication* between the two players, but in *presence of shared resources*. The ability of the two players to solve the game then depends on the type of resources they share. It is known that, for the so-called CHSH game, i.e., for the game $a \oplus b = x \wedge y$, the ability for the players to use entangled quantum bits (qubits) helps. We show that, apart from the CHSH game, quantum correlations do not help, in the sense that, for every game not equivalent to the CHSH game, there exists a classical protocol (using shared randomness) whose probability of success is at least as large as the one of any protocol using quantum resources. This result holds for both worst case and average case analysis. It is achieved by considering a model stronger than quantum correlations, the *non-signaling model*, which subsumes quantum mechanics, but is far easier to handle.

1 Introduction

One of the most celebrated results in the context of network computing is Linial's $\Omega(\log^* n)$ lower bound [32] on the number of rounds required for 3-coloring the n-node ring distributedly. In essence, this lower bound states that even if nodes can communicate an arbitrarily large amount of data between neighbors at every communication round, and even if nodes can perform an arbitrarily

[1] A preliminary version of this paper appeared in the proceedings of the 19th Int. Colloquium on Structural Information and Communication Complexity (SIROCCO), June 2012. Both authors are partially supported by the ANR project DISPLEXITY, and by the INRIA project GANG. Part of this work was done within the BQR project QDC granted by University Paris Diderot.

large amount of computation between every two communication rounds, 3-coloring the n-node ring in a distributed manner requires $\Omega(\log^* n)$ communication rounds. In other words, 3-coloring the ring requires some information to flow between nodes at distance larger than any constant. This lower bound is very robust. In particular, it holds even for Las Vegas algorithms where the nodes have access to a shared source of randomness.

In this paper, we question the *universality* of results such as Linial's lower bound. A lower bound (or an impossibility result) established for a distributed computing model offering very specific features has indeed little *conceptual* interest. (It may however have a significant *practical* interest if the model reflects a widely used technology). Instead, if the model is generic enough to capture a large number of frameworks, then the lower bound is quite significant conceptually. This is the case of Linial's lower bound, but up to some extent only. Indeed, on the one hand, the formal distributed computing model used in [32] – the $\mathcal{LOCAL}$ model – is very liberal, and therefore the lower bound remains valid in very many contexts. Still, the $\mathcal{LOCAL}$ model is based on classical physics, while there are several physical evidences[1] indicating that we may not be living in a world governed by classical physics. A natural question is therefore to ask whether, for example, the $\Omega(\log^* n)$ lower bound for 3-coloring the ring still holds if nodes are able to store, manipulate, and exchange resources such as, e.g., quantum bits (qubits).

At this point, we want to point out that the question of whether quantum computers able to manipulate a large number of qubits will one day exist is still open, and this paper is not aiming at arguing in favor or against this existence. Nevertheless, the practical efficiency of quantum effects in the context of distributed computing has already been demonstrated. One preeminent example is the establishment of long-distance quantum cryptographic channels between two parties[2]. Hence, while it is not completely clear whether connecting a large number of powerful quantum computers able to exchange very many qubits is achievable, it is a fact that the presence in a network of a handful of computers capable of manipulating just a few qubits may radically change the computational power of the network [10, 12, 16]. It is therefore of the utmost importance to determine whether or not the known limitations of network computing can be overcome by using quantum mechanic effects, and if so, to which extent.

Now, it is worth noticing that tackling the above question may not necessarily require any knowledge regarding quantum physics and/or quantum computing. One purpose of this paper is in fact to point out to the reader that, for several frameworks, and for several problems, lower bounds for quantum distributed computing can be derived *without manipulating any concepts related to quantum mechanics*. Indeed, there exist models that offer the same flavor as classical models, but subsume quantum distributed computing. Here is why: roughly, (classical) distributed algorithms using shared randomness produce outputs that are statistically distributed according to some specific kinds of distributions, and the same holds for the outputs produced by quantum distributed algorithms. Let us denote by $\mathcal{D}_{random}$ and $\mathcal{D}_{quantum}$ the set of distributions produced by the former and the latter, respectively. We have $\mathcal{D}_{random} \subset \mathcal{D}_{quantum}$, but the structure of $\mathcal{D}_{quantum}$ is quite difficult to handle for one not familiar with quantum computing. The good news is that there is a larger set of distributions, that is easier to handle than $\mathcal{D}_{quantum}$, but restricted enough so that non-trivial lower bounds can be designed for it. This latter set of distributions is called *non-signaling*, and is denoted here by $\mathcal{D}_{nonsignaling}$.

[1] John F. Clauser, Alain Aspect, and Anton Zeilinger received the Wolf Prize in Physics in 2010 for, in particular, their increasingly sophisticated series of tests of Bells inequalities.

[2] There are currently companies offering commercial quantum key distribution systems.

Informally, as the distributions in the sets $\mathcal{D}_{random}$ and $\mathcal{D}_{quantum}$ are constrained by classical and quantum mechanics, respectively, the distributions in the set $D_{nonsignaling}$ are only constrained by *relativistic causality*. The assumption of relativistic causality states that effects belong to the light cone of their causes, or, alternatively, *causal influences* do not travel faster than the speed of light. Hence, this type of distributions captures not only classical distributions, but also distributions that appear due to quantum effects. The reason why quantum correlations belong to the set $D_{nonsignaling}$ of non-signaling correlations may seem unclear to the reader unfamiliar with quantum computing. Indeed, some quantum phenomena seem to be in contradiction with relativistic causality. A typical case of such an apparent contradiction shows up when considering a system of a pair of infinitely distant particles that have been pre-set in entangled states, for which the measure of one particle's state gives immediate knowledge about the other particle's state. This phenomenon, which violates the classical concept of *local realism*, is known as quantum *nonlocality*. However, it is crucial to note that this phenomenon does not violate relativistic causality. Indeed, quantum nonlocality is just the expression of some particular probability distributions. As an example, the EPR paradox [19] involves two particles a and b, each one having a *spin* $+1$ or -1, such that the joint probability distribution of the pair of spins satisfies:

$$\begin{cases} \Pr[a=+1,\ b=-1] &=& \Pr[a=-1,\ b=+1] &=& \frac{1}{2} \\ \Pr[a=+1,\ b=+1] &=& \Pr[a=-1,\ b=-1] &=& 0. \end{cases}$$

In this case, a measurement performed on a gives immediate information on b, although no signals propagate from b to a. However, the marginal distributions of a and b are independent from each other (both are uniform in $\{-1,+1\}$). Prior to measurement, both have equal probabilities of being 0 or 1. The fact that the marginal distributions are independent of each other is the evidence that the above joint distribution is non-signaling. More generally, it has been proved that quantum correlations are nonlocal, but cannot be of any use for transmitting signals faster than light [28], and thus quantum correlations are non-signaling.

Non-signaling distributions form a stronger model than quantum computing. It is however unlikely to be a realistic model as argued in [15]. (If every non-signaling distribution could be realized in a physical experiment, then non plausible distributed computations could be achieved, like computing the scalar product of two vectors located at remote places by exchanging only one bit of information). Nevertheless, this paper is interested in impossibility results, and, as we will show, the known chain of strict containments

$$D_{random} \subset D_{quantum} \subset D_{nonsignaling}$$

enables to establish lower bounds on the power of distributed quantum computing, for several problems at least.

1.1 Our results

The main objective of this paper is to provide illustrative examples of distributed computing problems for which it is possible to design tight lower bounds for *quantum* algorithms without having to manipulate concepts from quantum mechanics, at all. This is achieved by considering the *non-signaling model*, that is stronger than distributed quantum computing, but involves no concepts from quantum mechanics.

As a case study, we address the following class of 2-player problems (see Figure 1). Alice (resp., Bob) receives a boolean x (resp., y) as input, and must return a boolean a (resp., b) as output. A

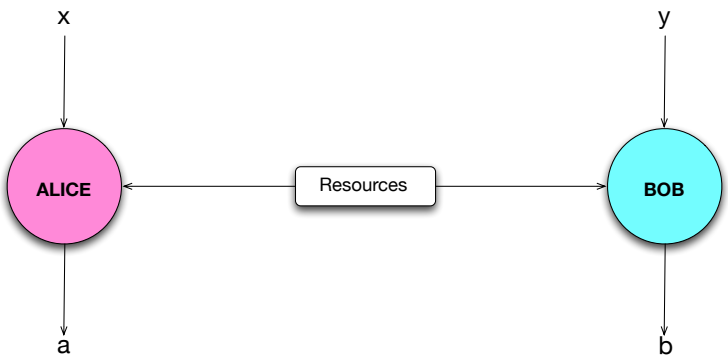

Figure 1: *Alice receives boolean x as input, while Bob receives boolean y. Alice and Bob are separated and cannot communicate. However, they had access to a common source of resources (e.g., random bits, intricate qubits, etc.) before being separated, and before getting their inputs. In the (δ, f) game, they have to compute boolean outputs a and b, respectively, such that $\delta(a, b) = f(x, y)$.*

game between Alice and Bob is defined by a pair (δ, f) of boolean functions. The objective of Alice and Bob playing game (δ, f) is, for every pair (x, y) of inputs, to output values a and b, respectively, satisfying

$$\delta(a, b) = f(x, y)$$

in *absence of any communication* between the two players. However, the two players have access to common resources such as a source of random bits, or a source of entangled quantum bits (qubits). The ability of solving a given game (δ, f) thus depends on the type of resources shared by Alice and Bob. It is known [14] that, for the so-called CHSH game

$$a \oplus b = x \wedge y \ ,$$

the ability for the players to use entangled qubits helps. We show that, apart from the CHSH game, and games equivalent to the CHSH game[3], quantum correlations do not help. That is, for every game non equivalent to the CHSH game, there exists a classical protocol (using shared randomness) whose probability of success is at least as large as the one of any protocol using quantum resources. This result holds for both worst case and average case analysis.

1.2 Related work

The paper [27] inspired us very much. In that paper, the authors define different extensions of the $\mathcal{LOCAL}$ model, by enabling processors to manipulate qubits, and they establish several separation results between these extensions. They also point out that lower bounds for distributed quantum computing can be obtained by considering a stronger model, called φ-$\mathcal{LOCAL}$ in [27]. It turns out that this latter model is noting else than the non-signaling model considered in this paper. We

[3]E.g., the game $a \oplus b = \bar{x} \wedge \bar{y}$ is equivalent to the CHSH game, as Alice and Bob just have to complement their respective inputs, and solve the CHSH game on these complemented inputs.

prefer using the terminology "non-signaling" because it is the standard terminology used in the physics literature.

There is a huge literature on design and analysis of algorithms in the $\mathcal{LOCAL}$ model of distributed computing. We refer to the book [36] for a presentation of some of the main achievements in this framework. As far as distributed graph coloring is concerned, we already mentioned the lower bound $\Omega(\log^* n)$ on the number of rounds for $(\Delta+1)$-coloring of n-node graphs [32], where Δ denotes the maximum degree. So far, the best known upper bound for deterministic algorithms is $2^{O(\sqrt{\log n})}$ rounds in [35], while the best known upper bound for randomized (Las vegas) algorithms is $O(\log n)$ expected number of rounds [1, 33]. These bounds can be improved for bounded degree graphs. Specifically, $(\Delta+1)$-coloring can be randomly computed in expected $O(\log \Delta + \sqrt{\log n})$ communication rounds (see [38]) recently improved to $O(\log \Delta + e^{O(\sqrt{\log \log n})})$ rounds in [5]. In contrast, the best known deterministic algorithm performs in $O(\Delta + \log^* n)$ rounds [4, 30].

The same way Monsieur Jourdain has been speaking prose without knowing it, many lower bounds in the literature are using arguments that can be directly applied to distributed quantum algorithms. As mentioned in [21], this is typically the case of "limited sight" arguments, since quantum mechanics respect causality. For instance, [27] noticed that the lower bounds for Maximal Independent Set [31], Locally-Minimal Coloring [26], and Sparse Connected Subgraph [17, 20] also hold for quantum computing. However, proofs based on other kinds of arguments, like, typically, Linial's lower bound [32], do not extend trivially to quantum computing models. (We shall come back to this issue later in the text).

Regarding distributed quantum computing, one can already cite a few contributions (see, e.g., the survey [10, 12]). For instance [16] lists a series of papers aiming at solving leader election in several variants of distributed quantum computing. Recently, [21] tackles several network problems (connectivity, MST, etc.) in a quantum computing model extending the $\mathcal{CONGEST}$ model [36] to capture quantum effects. Closer to our work are all contributions to multi-player "pseudo-telepathy" games (i.e., games solvable in absence of communications, using entanglement). We refer to the survey [9] for the analysis of several such games. In particular, the fact that the CHSH game [14] can be solved with probability $\cos^2(\pi/8)$ has been established in [8, 11], while [13] showed that this is the best success probability that can be achieved by a quantum strategy. Two-player games have also been investigated using the concept of boxes. A box is a conceptual device which receives pairs of inputs, and returns pairs of outputs distributed according to some non-signaling probability distribution which characterizes the box. The PR-box, solving the CHSH game with probability 1, has been introduced in [37], and is further studied in [6, 7]. It is known [7] that PR-boxes can be used to simulate any binary-output games. However, they cannot alone simulate any multiple-output games [18].

2 Non-signaling computation

Let us consider a distributed algorithm $\mathcal{A}$ performing in networks modeled as simple connected undirected graphs, under a synchronous message passing model [36]. In this model, processors have pairwise distinct identities. They are woken up simultaneously, and computation proceeds in fault-free synchronous rounds during which every processor exchanges messages of unlimited size with its neighbors, and performs arbitrary computations on its data.

More specifically, the n nodes of network $G = (V, E)$ are given pairwise distinct identities in $[n] = \{1, \ldots, n\}$, and we denote by $\mathrm{id}(u)$ the identity of node u. Let $\mathbf{x}$ be an n-dimensional vector

denoting the inputs to the nodes, where $\mathbf{x}_i \in \{0,1\}^*$ denotes the input binary string given to the node with identity i, for $i = 1, \ldots, n$. Similarly, let us denote by $\mathbf{y}_i \in \{0,1\}^*$ the output of the node with identity i, and let $\mathbf{y} = (\mathbf{y}_i)_{i\in[n]}$. Let $t = t(G, \mathrm{id}, \mathbf{x})$ be the running time of $\mathcal{A}$ for the input configuration $(G, \mathrm{id}, \mathbf{x})$ where G is an n-node network with nodes identified by $\mathrm{id} : V \to [n]$, and where the node with identity i receives input $\mathbf{x}_i$, for every $i \in [n]$. That is, the output $\mathbf{y}_i$ of the node with identity i is computed by $\mathcal{A}$ based solely on the ball $B_{G,\mathrm{id},\mathbf{x}}(i,t)$ of radius t in G, centered at node labeled i, including the structure of the subgraph of G induced by all nodes at distance at most t from i, together with the inputs and the identities of these nodes. In the sequel, when it is clear from the context, the subscripts G, id, and $\mathbf{x}$ will omitted, and $B_{G,\mathrm{id},\mathbf{x}}(i,t)$ simply denoted by $B(i,t)$.

A *task* T is defined by an input-output relation that, given a triple $(G, \mathrm{id}, \mathbf{x})$, specifies all valid output vectors $\mathbf{y}$ for graph G with nodes identified by id, and input $\mathbf{x}$. For instance, in the case of the k-coloring task, given $(G, \mathrm{id}, \mathbf{x})$, one must have (1) $\mathbf{y}_i \in \{1, \ldots, k\}$ for every $i \in [n]$, and (2) $\mathbf{y}_i \neq \mathbf{y}_j$ whenever the two nodes with respective identities i and j are adjacent in G.

We are interested in the different resources that can be accessed by an algorithm solving a task T in the above synchronous message passing model.

2.1 Local computing

If $\mathcal{A}$ is deterministic, then $\mathcal{A}$ is simply a mapping from D to $\{0,1\}^*$, where D is the *domain* of $\mathcal{A}$, that is, D is the set of all possible balls $B(i,t)$ that can be formed by legal inputs to $\mathcal{A}$. (E.g., if $\mathcal{A}$ is designed for planar graphs with boolean inputs given to nodes, then a ball $B(i,t)$ is legal if and only if it is planar, with $\mathbf{x}_j \in \{0,1\}$ for every node j in this ball). Therefore, the output vector $\mathbf{y} = (\mathbf{y}_i)_{i\in[n]}$ for some input configuration $(G, \mathrm{id}, \mathbf{x})$ is simply defined by its n coordinates $\mathbf{y}_i = \mathcal{A}(B(i,t))$ for $i = 1, \ldots, n$. This can be denoted as:

$$\mathbf{y} \mid (G, \mathrm{id}, \mathbf{x}) = \Big(\mathcal{A}(B(i,t)\Big)_{i\in[n]} \ . \tag{1}$$

Randomization provides additional power to the algorithm. In particular, assuming that every node is given a private source of random values ω in some probabilistic space Ω, then, for every $y \in \{0,1\}^*$, the probability that the node with identity i output y depends on $B(i,t)$ as well as on the collection of random values at the nodes in this ball. In this context, algorithm $\mathcal{A}$ yields a probability distribution on the possible outputs $\mathbf{y}$. That is, the output vector is now a random variable $\mathbf{Y}$, and

$$\mathbf{Y} \mid (G, \mathrm{id}, \mathbf{x}) = \Big(\mathbf{Y}_i \mid B(i,t)\Big)_{i\in[n]} \tag{2}$$

where $\mathbf{Y}_i$ is the ith coordinate of $\mathbf{Y}$, that is the random variable corresponding to the output at node i. Typically, if the nodes do not exchange their private random values, and if every node i acts by, first, collecting the data in $B(i,t)$, and, second, computing its (random) output $\mathbf{y}_i$ based on these data and on its private random coins, then the distribution of the output $\mathbf{Y}$ can be expressed, for any fixed $\mathbf{y} = (\mathbf{y}_i)_{i\in[n]}$, as:

$$\Pr[\mathbf{Y} = \mathbf{y} \mid (G, \mathrm{id}, \mathbf{x})] = \prod_{i=1}^{n} \Pr[\mathbf{Y}_i = \mathbf{y}_i \mid B(i,t)] \ .$$

In the above expression, $\Pr[\mathbf{y}_i \mid B(i,t)]$ denotes the distribution of the output $\mathbf{y}_i$ at node i applying Algorithm $\mathcal{A}$, knowing $B(i,t)$. This distribution depends in turn on the distribution of the private random values $\omega \in \Omega$ used at node i.

Obviously distributions that can be described by Eq. (2) subsume those that can be defined by Eq. (1). Shared randomness enlarges the spectrum of possible distributions even further. In the context of shared randomness, nodes have collectively access to a common source of random values λ in some probabilistic space Λ, in addition to possible private sources of randomness. The distribution of the output $\mathbf{Y}$ can then be expressed as:

$$\mathbf{Y} \mid (G, \mathrm{id}, \mathbf{x}) = \Big(\mathbf{Y}_i \mid B(i,t) \wedge \lambda\Big)_{i \in [n]} \text{ with probability } \Pr[\lambda], \text{ for every } \lambda \in \Lambda. \tag{3}$$

In the above expression, $\Pr[\lambda]$ denotes the distribution of the random value $\lambda \in \Lambda$, while $\mathbf{Y}_i \mid B(i,t) \wedge \lambda$ denotes the distribution of the output at node i applying Algorithm $\mathcal{A}$, knowing $B(i,t)$ and λ, which may also depend on the distribution of private values $\omega \in \Omega$. Again, obviously, the above expression subsumes the one in Eq. (2). Note also, that if the nodes proceed in first collecting all data in their ball of radius t, and then using their private values to compute their outputs, then

$$\Pr[\mathbf{Y} = \mathbf{y} \mid (G, \mathrm{id}, \mathbf{x})] = \sum_{\lambda \in \Lambda} \prod_{i=1}^{n} \Pr[\mathbf{Y}_i = \mathbf{y}_i \mid B(i,t) \wedge \lambda] \cdot \Pr[\lambda] \ .$$

All the previous expressions for $\mathbf{Y}$ in Equations (1), (2), and (3) capture locality in the sense that distant events can only be weakly correlated. That is, the outputs at two nodes i and j such that $B(i,t) \cap B(j,t) = \emptyset$ are only correlated according to the distributions generated by shared randomness. Nevertheless, it is worth pointing out that causality is not captured in its full generality by these equations. Indeed, roughly speaking, causality expresses the fact that the *distribution* of outputs at node i should not depend on events taking place at far away nodes, i.e., at nodes lying in G at distance greater than t. The expression of $\mathbf{Y}$ yielded by shared randomness satisfies this constraint. However, *there are distributions that satisfy causality which cannot be expressed as Eq.* (3). To see why, we need to formally define the notion of *non signaling* distributions, as introduced in the next section.

2.2 Non-signaling distributions

Let us consider a probability distribution on the output vectors $\mathbf{y} = (\mathbf{y}_i)_{i \in [n]}$, $\mathbf{y}_i \in \{0,1\}^*$, described by a random variable $\mathbf{Y}$ conditioned over all n-node graphs with nodes labeled with pairwise distinct identities in $[n]$, where the node with identity i is given input $\mathbf{x}_i \in \{0,1\}^*$, for every $i \in [n]$. More precisely, $\mathbf{Y}$ is defined conditionally to all possible triples $(G, \mathrm{id}, \mathbf{x})$ where id is the identity function for G, and $\mathbf{x} = (\mathbf{x}_i)_{i \in [n]}$, $\mathbf{x}_i \in \{0,1\}^*$, denotes the inputs given to the nodes. To describe $\mathbf{Y}$, we are thus given a collection of conditional distributions

$$\{\mathbf{Y} \mid (G, \mathrm{id}, \mathbf{x}) \ \text{ for all triples } (G, \mathrm{id}, \mathbf{x})\}. \tag{4}$$

where $\sum_{\mathbf{y}} \Pr[\mathbf{Y} = \mathbf{y} \mid (G, \mathrm{id}, \mathbf{x})] = 1$ for every configuration $(G, \mathrm{id}, \mathbf{x})$. Typical examples of such a distribution are those distributions as in Eq. (1), (2), and (3). For a fixed configuration $(G, \mathrm{id}, \mathbf{x})$, the distribution $\mathbf{Y} \mid (G, \mathrm{id}, \mathbf{x})$ yields the marginal distributions $\mathbf{Y}_i \mid (G, \mathrm{id}, \mathbf{x})$ defined over all binary strings in $\{0,1\}^*$, $i = 1, \ldots, n$. This distribution is corresponding to the way the i-th coordinate

of the output $\mathbf{Y}$ distributes whenever the whole vector $\mathbf{Y}$ distributes according to Eq. (4). More generally, for every subset $I \subseteq [n]$, the distribution $\mathbf{Y} \mid (G, \mathrm{id}, \mathbf{x})$ yields the marginal distributions

$$\mathbf{Y}_I \mid (G, \mathrm{id}, \mathbf{x})$$

defined over all $|I|$-dimensional vectors with coordinates in $\{0,1\}^*$. It is the distribution corresponding to the way the vector $\mathbf{Y}_I = (\mathbf{Y}_i)_{i \in I}$ distributes whenever the whole vector $\mathbf{Y}$ distributes according to Eq. (4). The set of marginal conditional distributions

$$\{\mathbf{Y}_I \mid (G, \mathrm{id}, \mathbf{x}), \text{ for all triples } (G, \mathrm{id}, \mathbf{x}) \text{ and all } I \subseteq [n]\}$$

enables to characterize whether or not the distribution given by Eq. (4) transmits "signals" at distance larger than t. Informally, if the set of balls of radius t, centered at nodes with identities in I, is the same for two configurations $(G, \mathrm{id}, \mathbf{x})$ and $(G', \mathrm{id}', \mathbf{x}')$, then the marginal distributions of the outputs of the nodes in I should be identical in both configurations. This is formally captured by the following definition, derived from [6, 27]. For any two random variables U and U', let $U \sim U'$ denotes the fact that U and U' are identically distributed.

Definition 1. *Let $t \geq 0$. For every positive integer n, a distribution $\mathbf{Y}$ described by its conditional distributions as in Eq.* (4) *is* non-signaling *at distance more than t if, for every two n-node graphs G and G' with respective identity assignments* id *and* id$'$*, and respective inputs $\mathbf{x}$ and $\mathbf{x}'$ given to their nodes, and for every $I \subseteq [n]$, the marginal distribution $\mathbf{Y}_I$ satisfy the following:*

$$B_{G,\mathrm{id},\mathbf{x}}(i,t) = B_{G',\mathrm{id}',\mathbf{x}'}(i,t) \text{ for all } i \in I \implies \mathbf{Y}_I \mid (G, \mathrm{id}, \mathbf{x}) \sim \mathbf{Y}_I \mid (G', \mathrm{id}', \mathbf{x}'). \tag{5}$$

In other words, to be non-signaling at distance more than t, the output distribution $\mathbf{Y}$ must satisfy that if $B_{G,\mathrm{id},\mathbf{x}}(i,t) = B_{G',\mathrm{id}',\mathbf{x}'}(i,t)$ for every $i \in I$, then, for every $|I|$-dimensional vector $\mathbf{z}$ with coordinate in $\{0,1\}^*$, we must have

$$\Pr[\mathbf{Y}_I = \mathbf{z} \mid (G, \mathrm{id}, \mathbf{x})] = \Pr[\mathbf{Y}_I = \mathbf{z} \mid (G', \mathrm{id}', \mathbf{x}')].$$

In particular, from the definition above, if $B_{G,\mathrm{id},x}(i,t) = B_{G',\mathrm{id}',x'}(i,t)$ for some $i \in [n]$, then the non-signaling condition states that the distributions of the outputs at node i must be identical in $(G, \mathrm{id}, \mathbf{x})$ and $(G', \mathrm{id}', \mathbf{x}')$. The non-signaling condition is actually more constrained by requesting that the property should hold not only at every individual node, but also for all possible scales of computation, that is, for every subset $I \subseteq [n]$, to capture the case of pairwise independent events that are not mutually independent. The condition $B_{G,\mathrm{id},\mathbf{x}}(i,t) = B_{G',\mathrm{id}',\mathbf{x}'}(i,t)$ for every $i \in I$ states that, for every node with identity $i \in I$, all the information this node can gather in t rounds is identical in both instances $(G, \mathrm{id}, \mathbf{x})$ and $(G', \mathrm{id}', \mathbf{x}')$. The non-signaling condition states that whenever this is the case, the marginal distributions of the outputs at the nodes with identities in I must be identical in both instances $(G, \mathrm{id}, \mathbf{x})$ and $(G', \mathrm{id}', \mathbf{x}')$, which is expressed in Eq. (5).

Remark. Consider a distribution $\mathbf{Y}$ that violates Eq. (5). It means that there exists a set I, and two distinct instances $(G, \mathrm{id}, \mathbf{x})$ and $(G', \mathrm{id}', \mathbf{x}')$ such that $B_{G,\mathrm{id},\mathbf{x}}(i,t) = B_{G',\mathrm{id}',\mathbf{x}'}(i,t)$ for all $i \in I$ while $\mathbf{Y}_I \mid (G, \mathrm{id}, \mathbf{x})$ distributes differently from $\mathbf{Y}_I \mid (G', \mathrm{id}', \mathbf{x}')$. In other words, the "behavior" of the nodes with identities in I differ in $(G, \mathrm{id}, \mathbf{x})$ and $(G', \mathrm{id}', \mathbf{x}')$, as witnessed by the fact that the output vector $\mathbf{Y}_I$ is not distributed the same in both instances, whereas the "view" at distance t of the nodes in I are identical in both instances. As a consequence, for such a distribution, it means that some "signal" from nodes at distance greater than t reaches some nodes with identities in I, in no more than t rounds. This is not possible, unless the distribution violates causality.

2.3 Non-classical computing

By definition, every output distribution resulting from the execution of a distributed algorithm using shared randomness, and performing in at most t rounds, is non-signaling at distance greater than t. This is because, as illustrated by Eq. (3), the output of every node i is precisely conditioned on $B(i,t)$, and on the value of the shared random variable λ, and the output $\mathbf{Y}$ is simply the joint distribution of the marginal outputs $\mathbf{Y}_i$. Hence, if $B_{G,\mathrm{id},\mathbf{x}}(i,t) = B_{G',\mathrm{id}',\mathbf{x}'}(i,t)$ for all $i \in I$, then we get the desired non-signaling equality $\mathbf{Y}_I \mid (G,\mathrm{id},\mathbf{x}) \sim \mathbf{Y}_I \mid (G',\mathrm{id}',\mathbf{x}')$.

Bell's inequalities (cf. the next section) precisely states that no *classical* distributed algorithms (i.e., algorithms based on classical physics) can yield all distributions to be observed in quantum mechanics [8], and such distributions have been experimentally observed (see [25] for the first experimental test of the violation of Bell's inequalities), demonstrating that we may not live in a world governed by classical physics. In particular, it is known that quantum effects generate correlations enabling to go beyond the distribution of Eq. (3). A very basic example explored exhaustively in this paper is the scenario in which two processes Alice and Bob have access to some variables u and v, respectively, correlated in a way consistent with quantum physics. This correlation may not be captured by shared randomness. Thus, an algorithm run by Alice and Bob with access to the variable u at A, and v at B, might be able to solve tasks quicker, or with higher success probability than when the algorithm is restricted to the access to private and/or shared random variables only. How to exploit the power of probabilistic correlations beyond the ones resulting from using shared randomness is the the purpose of *quantum distributed computing*.

Quantum distributed computing is rather in its infancy, but we have already mentioned a set of quite significant contributions. Designing efficient quantum distributed algorithms is a challenge, and requires a very fine and deep skill in the manipulation of quantum correlations. There are good news though: designing *lower bounds* for quantum distributed computing may not be as hard as designing algorithms, at least in some frameworks. More precisely, there are several tasks for which lower bounds or impossibility results designed for classical algorithms translate easily to the context of quantum computing. This statement of facts is thanks to non-signaling distributions, as defined in Definition 1.

Indeed, consider a distribution $\mathbf{Y}$ that is non-signaling at distance greater than t. Define the success probability of $\mathbf{Y}$ for a task T as follows. The success probability of $\mathbf{Y}$ for $(G,\mathrm{id},\mathbf{x})$ is the sum, over all *valid* outputs $\mathbf{y}$ for T with respect to the configuration $(G,\mathrm{id},\mathbf{x})$, of the probability of $\mathbf{y}$. That is:

$$\Pr[\mathbf{Y} \text{ succeeds for } (G,\mathrm{id},\mathbf{x})] = \sum_{\text{valid } \mathbf{y}} \Pr[\mathbf{Y} = \mathbf{y} \mid (G,\mathrm{id},\mathbf{x})].$$

Then, the success probability of $\mathbf{Y}$ for the task T is the infimum, taken over all configurations $(G,\mathrm{id},\mathbf{x})$, of the success probability of $\mathbf{Y}$ for $(G,\mathrm{id},\mathbf{x})$. Now, quantum physics does not violate causality, and all distributions generated by quantum distributed algorithms communicating at distance no more than t are non-signaling at distance greater than t. Hence, given a task T, proving that there are no distributions non-signaling at distance greater than t with success probability greater than p for T enables to prove that every distributed quantum algorithm solving T with probability at least p runs in at least t rounds.

The following result is folklore. It simply states that quantum mechanics does not violate causality.

Theorem 1 (see, e.g., [27])**.** *For any $t \geq 0$, any output distribution produced by a quantum dis-*

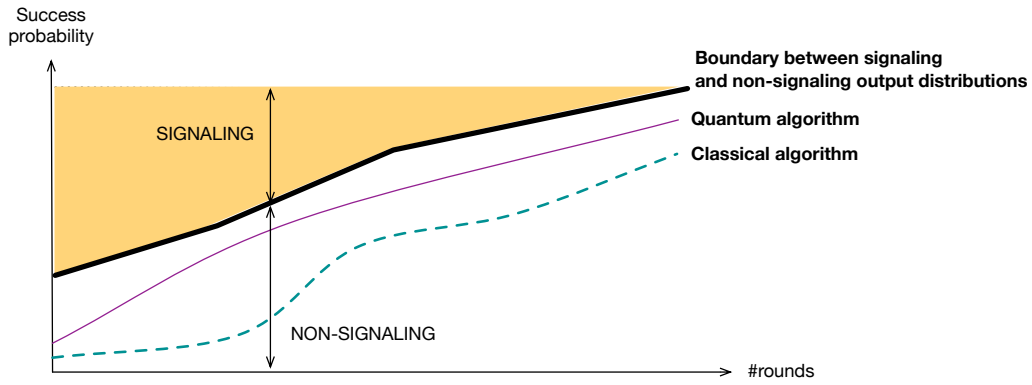

Figure 2: *Quantum resources enable to design distributed algorithms that are at least as efficient as any classical algorithm (even those using shared randomness), but the output distribution of any distributed quantum algorithm cannot offer better tradeoff between success probability and number of rounds than the best non-signaling distribution.*

tributed algorithm performing t communication rounds is non-signaling at distance greater than t.

As a consequence, we immediately get:

Corollary 1. *Let T be a task, and let $t \geq 0$. Assume that, for every distribution $\mathbf{Y}$ that is non-signaling at distance greater than t, the success probability of $\mathbf{Y}$ for T is at most p. Then, there are no quantum distributed algorithms enabling to solve task T with probability more than p in t communication rounds.*

Figure 2 displays an abstract graphical representation of the tradeoff between time complexity and success probability for various kinds of algorithms. We stress the fact that the world "quantum" in the statement of Corollary 1 can be replaced by any distributed computing model M, whether it be weaker, stronger, orthogonal, or inconsistent with quantum computing, as long as the model M is non-signaling.

2.4 Examples and applications

The following example is borrowed from [27], by the courtesy of the authors: Cyril Gavoille, Adrian Kosowski, and Marcin Markiewicz. Consider the task of 2-coloring the nodes in a ring C_n with an even number n of nodes. Linial [32] has proved that, with a classical algorithm, $n/2 - 1$ rounds are necessary and sufficient for solving that task with probability 1. (The lower bound holds even for an algorithm using shared randomness, while the upper bound holds deterministically). Let us show that, as pointed out in [27], quantum algorithms cannot do much better in term of number of rounds, in the sense that at least $n/4-1$ rounds are required by such algorithms for 2-coloring rings with an even number of nodes. Moreover, this holds even if one just asks for a success probability greater than $1/2$. To establish this claim, we use Corollary 1, and prove that any distribution $\mathbf{Y}$ that is non-signaling at distance greater that $n/4 - 1$ cannot solve 2-coloring with probability greater than $1/2$.

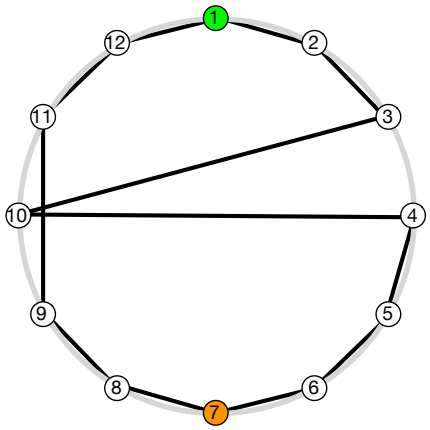

Figure 3: *Two rings on 12 nodes (in black, and in light grey). For $t = 2$, the balls $B(1, t)$ and the balls $B(7, t)$ are identical in both rings. In a proper 2-coloring, nodes 1 and 7 must have different colors in one ring, while they must have identical color in the other ring.*

Let $n = 4(t + 1)$ for $t \geq 1$. Hence, two antipodal nodes in the n-node ring C_n (at distance $n/2$) satisfies that the balls of radius t centered at these nodes do not intersect, and are in fact separated by two nodes which belong to none of these two balls. (See Figure 3). We consider two configurations $(C_n, \mathrm{id}, \varnothing)$ and $(C_n, \mathrm{id}', \varnothing)$ (the coloring problem assumes no inputs). In $(C_n, \mathrm{id}, \varnothing)$, the nodes of the rings are labeled consecutively from 1 to n. In $(C_n, \mathrm{id}', \varnothing)$, nodes are also labeled consecutively from 1 to n, apart from one identity, $3(t + 1) + 1$, which is placed between $t + 1$ and $t + 2$, and identities $3(t+1)$ and $3(t+1)+2$ which are adjacent. Now, in the two configurations, the balls $B(1, t)$ are identical. Similarly, the balls $B(n/2 + 1, t)$ are identical in the two configurations. However, nodes 1 and $n/2 + 1$ are at even distance $2(t + 1)$ in $(C_n, \mathrm{id}, \varnothing)$ while they are at odd distance $2t + 1$ in $(C_n, \mathrm{id}', \varnothing)$.

Let $\mathbf{Y}$ be a non-signaling distribution with success probability p. It must be the case that, with probability at least p, nodes 1 and $n/2 + 1$ are colored the same in $(C_n, \mathrm{id}, \varnothing)$, but with different colors in $(C_n, \mathrm{id}', \varnothing)$. So, let $I = \{1, n/2 + 1\}$, and let us focus on $\mathbf{Y}_I$ conditioned on the two configurations $(C_n, \mathrm{id}, \varnothing)$ and $(C_n, \mathrm{id}', \varnothing)$. These distributions act on 2-dimensional boolean vectors $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2)$. Because of the way $\mathbf{z}$ is distributed according to $\mathbf{Y}_I \mid (C_n, \mathrm{id}, \varnothing)$, we must have $\Pr[\mathbf{z}_1 = \mathbf{z}_2] \geq p$ because 1 and $n/2+1$ are colored the same in $(C_n, \mathrm{id}, \varnothing)$, with probability at least p. Similarly, because of the way $\mathbf{z}$ is distributed according to $\mathbf{Y}_I \mid (C_n, \mathrm{id}', \varnothing)$, we must have $\Pr[\mathbf{z}_1 \neq \mathbf{z}_2] \geq p$.

Now, the non-signaling condition actually states that these two conditional distributions $\mathbf{Y}_I \mid (C_n, \mathrm{id}, \varnothing)$ and $\mathbf{Y}_I \mid (C_n, \mathrm{id}', \varnothing)$ are identical. Therefore, $1 - p \geq p$, i.e., $p \leq 1/2$. Thus, by Corollary 1, no quantum algorithms can successively 2-color the ring of even size with probability greater than $1/2$ in less than $n/4 - 1$ rounds. It is worth noticing that the above lower bound exploits Corollary 1 maximally, in the sense that there exists a distribution $\mathbf{Y}$ that is non-signaling at distance greater that $n/4$, with success probability 1 for 2-coloring rings with even number of nodes (see [27]).

In the remaining of this paper, we use the concepts introduced in this section, and specifically

Corollary 1, to exhaustively study 2-player games. These games involve two parties that do not communicate, but share resources (e.g., random bits, quantum bits, etc.).

3 Two-player games

We consider the following 2-player problem. Alice (resp., Bob) receives a boolean x (resp., y) as input, and must return a boolean a (resp., b) as output. A *game* between Alice and Bob is defined by a pair (δ, f) of boolean functions. The objective of Alice and Bob playing game (δ, f) is, for every inputs x and y, to output values a and b satisfying

$$\delta(a, b) = f(x, y)$$

in *absence of any communication* between the two players. Obviously, the game is trivial whenever there exist two boolean functions α and β such that $\delta(\alpha(x), \beta(y)) = f(x, y)$ for every pair $(x, y) \in \{0, 1\}^2$. Indeed, for such games, there exists a deterministic distributed protocol solving the game, with Alice returning $\alpha(x)$ on input x, and Bob returning $\beta(y)$ on input y. Non-trivial games may still be solved, but only under some probabilistic guarantees. A game (δ, f) is said to be solvable with probability p if there exists a (randomized) distributed protocol such that Alice outputs a, and Bob outputs b, with

$$\Pr(\delta(a, b) = f(x, y)) \geq p \tag{6}$$

for every input pair $(x, y) \in \{0, 1\}^2$.

Different sources of randomness can then be considered. Classical sources of randomness (i.e., not using quantum effects) include the case where each of the two players are provided with individual independent sources of random bits. They also include shared randomness where, in addition to individual independent sources of random bits, the two players have access to a common source of random bits. Rewriting Eq. (3) in the context of 2-player games, shared randomness enables to produce outputs satisfying

$$\Pr(a, b|x, y) = \sum_{\lambda \in \Lambda} \Pr(a|x, \lambda) \cdot \Pr(b|y, \lambda) \cdot \Pr(\lambda) \tag{7}$$

where the random variable λ is drawn from some probability space Λ, and $\Pr(a, b|x, y)$ denotes the probability that Alice outputs a and Bob outputs b, given the fact that Alice receives x as input, and Bob receives y as input. It is known [8] that correlations on quantum entangled states enable to derive protocols whose output distribution cannot be modeled as Eq. (7). One evidence of this fact is the CHSH game [14]:

$$a \oplus b = x \wedge y$$

where $\oplus$ denotes the exclusive-or operator. CHSH can be solved with probability $\cos^2(\pi/8) > \frac{3}{4}$ with a quantum protocol (i.e., a protocol in which Alice and Bob have access to entangled quits) [13], while every protocol using classical shared randomness cannot solve CHSH with probability more than $\frac{3}{4}$. In fact, the literature dealing with 2-player games [11] refers to objects called *boxes*. A box B is characterized by the probabilities $\Pr(a, b|x, y)$ of outputting pair (a, b) given the input pair (x, y), for all $a, b, x, y \in \{0, 1\}$. A box B is thus described by a set

$$\{\Pr(\cdot, \cdot|x, y),\ (x, y) \in \{0, 1\}^2\} \tag{8}$$

of four probability distributions on $\{0,1\}^2$, one for each pair $(x,y) \in \{0,1\}^2$. Hence, there are infinitely many boxes, with different computational powers.

The absence of communication between the two players along with the assumption of causality are captured by the class of *non-signaling* boxes, defined according to Definition 1. A box B is non-signaling if and only if it satisfies that the marginal output distributions for Alice and Bob depend only on their respective inputs. Formally, a non-signaling box satisfies:

$$\begin{aligned} &\forall a, x, \ \sum_b \Pr(a,b|x,0) = \sum_b \Pr(a,b|x,1), \\ \text{and} \quad &\forall b, y, \ \sum_a \Pr(a,b|0,y) = \sum_a \Pr(a,b|1,y). \end{aligned} \tag{9}$$

Non-signaling boxes satisfying Eq. (7) are called *local*, where "locality" is referring here to the physical science concept of *local hidden variables* [8, 19]. Boxes that do not satisfy Eq. (9) are *signaling*. As mentioned before, signaling boxes are not considered physically realistic because they would imply instantaneous transmission of signals between two distant entities.

The set of all boxes has a geometric interpretation [6], because it forms a 12-dimensional convex polytope corresponding to the four sets in Eq. 8, where each set is entirely described by three values (as the 4th is the sum of the first three). This polytope includes the convex polytope of non-signaling boxes, which includes in turn the convex local polytope. Fig. 4 provides an abstract representation of the non-signaling polytope. It is known [7] that each of the extremal vertices of the non-signaling polytope is equivalent (up to individual reversible transformations on the inputs and outputs) to the so-called PR box [13, 37], that is described by the distribution:

$$\Pr(a,b|x,y) = \begin{cases} \frac{1}{2} & \text{if } a \oplus b = x \wedge y \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the PR box satisfies $\Pr(a \oplus b = x \wedge y) = 1$ for every input pair x, y. So, in particular, it solves the CHSH game with probability 1. Each of the extremal vertices of the local polytope can be implemented by a deterministic protocol: they are "equivalent" to the identity box ID described by $\Pr(a,b|x,y) = 1$ if and only if $a = x$ and $y = b$. Every non-extremal box B is a linear combinations of extremal boxes: $B = \sum_{i=1}^{k} \beta_i B_i$ where B_i is an extremal box, $\sum_{i=1}^{k} \beta_i = 1$, and $\beta_i > 0$ for every $i = 1, \ldots, k$. In Fig. 4, the dotted line represents the limit of the class of boxes implementable by a quantum protocol. This latter class strictly contains the local boxes, and is strictly included in the class of non-signaling boxes, as witnessed by the CHSH game.

Our objective of identifying the games for which quantum correlations help can be reformulated as follows. Given a box implementable by a quantum protocol, which games can be efficiently solved using this box? Stated differently, given a game, what are the boxes implementable by a quantum protocol that enable to solve that game with better guarantees than any local boxes?

4 Equivalence classes of games

As introduced in the previous section, a game between Alice and Bob is described by a pair (δ, f) of boolean functions on two variables. Playing the game means for Alice (resp. Bob) to receive a boolean x (resp., y) as input, and to return a boolean a (resp., b) as output such that $\delta(a,b) = f(x,y)$ without communication between the two players. Examples of games are

$$\text{EQ}: \quad a \wedge b = \overline{x \oplus y} \quad \text{and} \quad \text{NEQ}: \quad a \wedge b = x \oplus y.$$

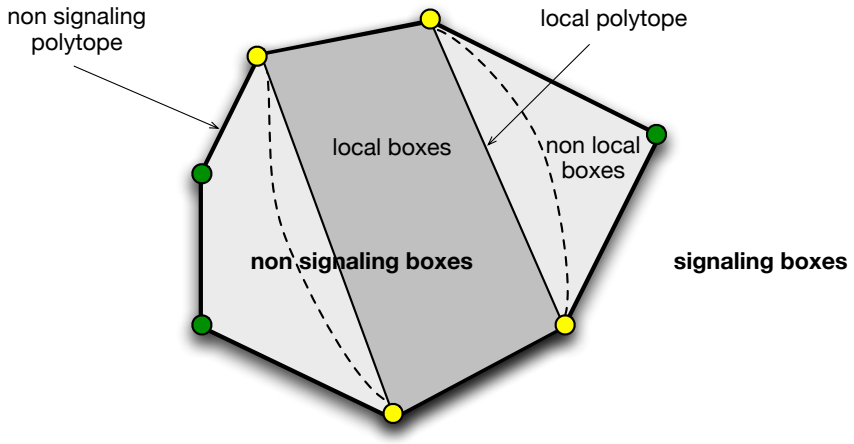

Figure 4: *Abstract representation of the non-signaling polytope, including the polytope of local boxes. The dotted line indicates the boundary between boxes which can be implemented by a quantum algorithm, and boxes which cannot.*

Another example of a game is :

$$\text{AMOS}: \quad a \wedge b = \overline{x \wedge y}.$$

In these three examples, one can view the games as Alice and Bob respectively deciding whether the equality $x = y$ holds, whether the non-equality $x \neq y$ holds, and whether there is "at most one selected" player (a selected player has input 1). Here, "deciding" means that if the answer is "yes" then both players should output "yes", while if the answer is "no" then at least one player should output "no". In fact, the three games EQ, NEQ, and AMOS, are AND-games, in the sense that δ is the conjunctive operator. However, all games are not of that type. In particular, we shall see that the already mentioned CHSH game

$$a \oplus b = x \wedge y$$

is not an AND-game, since it is not equivalent to any game (δ, f) where δ is the conjunctive operator. More precisely, for any game (δ, f), both functions δ and f can be rewritten as:

$$\delta(a,b) = \alpha_{1,1}ab + \alpha_{1,0}a + \alpha_{0,1}b + \alpha_{0,0} \text{ and } f(x,y) = \beta_{1,1}xy + \beta_{1,0}x + \beta_{0,1}y + \beta_{0,0}$$

where the $+$ symbol denotes the exclusive-or operator $\oplus$, the (omitted) $\cdot$ symbol denotes the and-operator $\wedge$, and all coefficients are in $\{0,1\}$. We say that two games (δ, f) and (δ', f') are equivalent if

$$\delta(a,b) = \delta'(A,B) \text{ and } f(x,y) = f'(X,Y)$$

where A (resp., B, X, Y) is a degree-1 polynomial in a (resp., b, x, y) with coefficients in $\{0,1\}$. Whenever two games are equivalent, any protocol solving one of the two games can be used for solving the other game, by performing individual reversible transformations on the inputs and outputs. The probability of success for the two games will be identical. The same notion of equivalence can be defined for boxes. Now we can state formally that the CHSH game is not equivalent to any of the three AND-games: EQ, NEQ, or AMOS. This is because, as we will see further

in the text, none of these latter games can be solved with probability 1 by a non-signaling box (as opposed to the CHSH game which can be solved with probability 1 by the PR box). Instead, EQ and NEQ are equivalent games. Indeed, for NEQ, $f(x,y) = x + y$, while, for EQ, $f(x,y) = x + (y+1)$.

Definition 2. *A game (δ, f) is an* XOR*-game if and only if it is equivalent to a game (δ', f') where $\delta'(a,b) = a \oplus b$.*

5 On the power of quantum correlations

In this section, we establish our main result, stating that correlations on quantum entangled states do not help for solving 2-player games that are not equivalent to an XOR-game. In fact we establish a stronger result by showing that non-signaling boxes do not help for those games.

Theorem 2. *Let (δ, f) be a 2-player game that is not equivalent to any* XOR*-game. Let p be the largest success probability for (δ, f) over all local boxes. Then every box solving (δ, f) with probabilistic guarantee $> p$ is signaling.*

Proof. The proof is straightforward for games (δ, f) where δ does not depend on both a and b. Indeed, on the one hand, if δ is constant, say $\delta(a,b) = \alpha$ for some $\alpha \in \{0,1\}$, for all (a,b), then the game is either impossible (whenever $\exists x, y : f(x,y) \neq \alpha$) or trivial (whenever $\forall x, y, f(x,y) = \alpha$). And, on the other hand, if δ depends on only one of its two variables, say $\delta(a,b) = a + \alpha$ for some $\alpha \in \{0,1\}$, then the game is again either trivial, or equivalent to a single-player game where the player must compute a 2-variable function $f(x,y)$ knowing only one of the variables. Games of that latter class are equivalent to either the game $a = y$ or the game $b = x$. Non-signaling boxes do not help for such games: the best probability of success is $\frac{1}{2}$ (which is achievable by a classical algorithm using randomization). Indeed, consider the game $a = y$, and, for that game, consider the instance $x = 0$ and $y = 0$. Then, a box B with success probability p satisfies that $\Pr[\text{success for input } (0,0)] \geq p$. Now,

$$\Pr[\text{success for input } (0,0)] = \sum_b \Pr(0,b|0,0),$$

and the non-signaling condition in Eq (9) states that $\sum_b \Pr(0,b|0,0) = \sum_b \Pr(0,b|0,1)$. The latter sum is the probability of failure for the input $(0,1)$, which is at most $1-p$. Therefore, $p \leq 1-p$, and thus $p \leq \frac{1}{2}$.

Therefore, we focus now on "true" 2-player games, i.e., games (δ, f) where δ depends on both a and b. First, we show that every true 2-player game (δ, f) which is not equivalent to an XOR-game is either deterministic, or equivalent to NEQ or AMOS. To establish this claim, observe that if f is constant, or depends on only one of the the two inputs, then the game (δ, f) can be solved with probability 1, by a deterministic protocol. Indeed, assume, without loss of generality, that f depends only on x. (The case f constant is straightforward). Then Alice and Bob can agree beforehand on a fixed value b^* for b. It follows that, knowing b^*, f, and δ, Alice can output a such that $\delta(a, b^*) = f(x)$.

We now come to the interesting case, that is, when both δ and f depend on their two inputs. Any 2-variable boolean function g can be rewritten as :

$$g(u,v) = U + V \quad \text{or} \quad g(u,v) = UV \quad \text{or} \quad g(u,v) = UV + 1$$

where U (resp., V) is a polynomial in u (resp., v) of degree at most 1, with coefficients in $\{0,1\}$. Given that fact, we rewrite any game (δ, f) using two expressions from the above, one for δ, and the other for f. We thus get nine different types of games, which can be narrowed down to five types by noticing that games like $A+B=XY+1$ are the same as games like $A'+B'=X'Y'$, up to the (reversible) transformation $B'=B+1$. These five types of games are the following:

$$\begin{aligned}
\delta(a,b)=A+B &= f(x,y)=X+Y\\
\delta(a,b)=A+B &= f(x,y)=XY\\
\delta(a,b)=AB &= f(x,y)=X+Y\\
\delta(a,b)=AB &= f(x,y)=XY\\
\delta(a,b)=AB &= f(x,y)=XY+1
\end{aligned}$$

Since f (resp., δ) depends on both x and y (resp., both a and b), all polynomials A, B, X, and Y in the above five types of games are of degree exactly 1, hence making all transformations reversible. Therefore, if two games can be rewritten into the same type, then they are equivalent. Table 1 describes the equivalence classes over the set of games formed by the five types above, and provides a representative for each class.

	Form of the class	Representative of the class
Deterministic	$AB=XY$	PROD $a\wedge b=x\wedge y$
	$A+B=X+Y$	SUM $a\oplus b=x\oplus y$
Not deterministic	$A+B=XY$	CHSH $a\oplus b=x\wedge y$
	$AB=X+Y$	NEQ $a\wedge b=x\oplus y$
	$AB=XY+1$	AMOS $a\wedge b=\neg(x\wedge y)$

Table 1: Equivalence classes for 2-player games depending on both inputs. The first two classes of games are deterministic, i.e., can be solved by a deterministic protocol. Instead, the last three classes are not deterministic (no deterministic protocols can solve any of the games in these three classes).

The theorem holds for games PROD and SUM since both of them can be solved by a deterministic protocol. Every game that is neither deterministic nor equivalent to an XOR-game is equivalent to the AND-game NEQ or AMOS. We now show that non-local boxes fail to solve AMOS or NEQ with higher probabilistic guarantee than what can be achieved with local boxes.

Let us first examine AMOS. We start by showing that any box that solves AMOS with probabilistic guarantee $p>\frac{2}{3}$ is signaling. Suppose that there exists a non-signaling box B, defined by the correlation $\Pr(a,b|x,y)$, that solves AMOS with probability p. On the one hand, for any probability distribution $\pi=\{\pi_{xy}|(x,y)\in\{0,1\}^2\}$ of the inputs, we have

$$\sum_{xy}\pi_{xy}\ \Pr(\text{success for input }(x,y))\geq p$$

On the other hand, we have

$$\sum_{xy}\pi_{xy}\Pr(\text{success for input }(x,y))=\sum_{xy}\pi_{xy}\sum_{ab}\mathbb{1}_{\{a\wedge b=\neg(x\wedge y)\}}\Pr(a,b|x,y)$$

where $\mathbb{1}_{\{a\wedge b=\neg(x\wedge y)\}}$ denotes the boolean indicator function of whether $a \wedge b = \neg(x \wedge y)$ is true or not. Let us consider the following distribution π^*:

$$\pi^*_{00} = 0 \quad \text{and} \quad \pi^*_{xy} = \frac{1}{3} \text{ for all } (x,y) \neq (0,0)$$

Let $p_{abxy} = \Pr(a,b|x,y)$ for box B. The probability of success with the input distribution π^* satisfies

$$\begin{aligned}
\sum_{xy} \pi^*_{xy} \Pr(\text{success for } (x,y)) &= \frac{1}{3} \sum_{xy\neq 00} \mathbb{1}_{\{a\wedge b=\neg(x\wedge y)\}} p_{abxy} \\
&= \frac{1}{3} \left(p_{1101} + p_{1110} + \sum_{ab\neq 11} p_{ab11}\right) \\
&= \frac{1}{3} \left(p_{1101} + p_{1110} + p_{0011} + p_{0111} + p_{1011}\right) \qquad (10)
\end{aligned}$$

The non-signaling conditions (cf., Eq. (9)) require that, for every a, b, x, y,

$$\begin{aligned}
p_{a0x0} + p_{a1x0} &= p_{a0x1} + p_{a1x1} \\
\text{and} \quad p_{0b0y} + p_{1b0y} &= p_{0b1y} + p_{1b1y}
\end{aligned}$$

This gives a bound on the first two terms of Eq. (10):

$$\begin{aligned}
p_{1101} &= p_{1111} + p_{0111} - p_{0101} \leq p_{1111} + p_{0111} \\
\text{and} \quad p_{1110} &= p_{1111} + p_{1011} - p_{1010} \leq p_{1111} + p_{1011}
\end{aligned}$$

The probability p of success is therefore bounded by :

$$\begin{aligned}
p &\leq \frac{1}{3} \left(p_{1111} + p_{0111} + p_{1011} + p_{1111} + p_{0011} + p_{0111} + p_{1011}\right) \\
&\leq \frac{1}{3} \Big(2 \sum_{ab} p_{ab11} - p_{1111}\Big) \\
&\leq \frac{2}{3}
\end{aligned}$$

Indeed, $\sum_{ab} p_{abxy} = 1$ for any fixed (x,y), and $p_{1111} \geq 0$. Therefore, every non-signaling box solves AMOS with success at most $\frac{2}{3}$.

Regarding NEQ, we observe that with distribution π^* that discards the input $(0,0)$, AMOS and NEQ become the same games:

$$f_{\text{AMOS}}(x,y) = f_{\text{NEQ}}(x,y)$$

for all $(x,y) \neq (0,0)$. As a consequence, the same bound $\frac{2}{3}$ also holds for NEQ: every non-signaling box solves NEQ with success at most $\frac{2}{3}$.

We now show that the bound $\frac{2}{3}$ for AMOS and NEQ can be reached by local boxes. For this purpose, we describe a protocol using solely shared randomness, and reaches success probability $\frac{2}{3}$. Let a_0 and a_1 (resp., b_0 and b_1) be the outputs of Alice (resp. Bob) on the respective input $x = 0$ and $x = 1$ (resp., $y = 0$ and $y = 1$). AMOS translates into solving the system:

$$\begin{cases}
a_0 \cdot b_0 = 1 \\
a_0 \cdot b_1 = 1 \\
a_1 \cdot b_0 = 1 \\
a_1 \cdot b_1 = 0
\end{cases} \qquad (11)$$

and NEQ translates into :

$$\begin{cases} a_0 \cdot b_0 = 0 \\ a_0 \cdot b_1 = 1 \\ a_1 \cdot b_0 = 1 \\ a_1 \cdot b_1 = 0 \end{cases} \tag{12}$$

The second and third equations of the system for AMOS as well as for NEQ imply that $a_0 = a_1 = b_0 = b_1 = 1$, resulting in the last equation impossible to be satisfied in both games. Hence the fourth equation of each system cannot be simultaneously satisfied with those two equations. Instead, if one chooses to ignore one of them, then one can find a solution to the game. Playing any one of the two games using shared randomness, we allow Alice and Bob to have access, before knowing their inputs, to a shared random variable λ uniformly distributed in $\{1, 2, 3\}$, designating the equation to be ignored among the last three ones. Alice and Bob will fail to solve the game with probability at most $\frac{1}{3}$ (when the ignored equation is precisely the one corresponding to the actual inputs), making the success probability for any input (x, y) equal to $\frac{2}{3}$. This completes the proof of the theorem. □

It turns out that even relaxing the constraints placed on solving the game, by considering average case analysis, does not allow non-signaling boxes to perform better than local boxes on games not equivalent to XOR-games.

Theorem 3. *Let (δ, f) be a 2-player game that is not equivalent to any* XOR*-game. Let p be the largest* average *success probability for (δ, f) over all local boxes. Then every box solving (δ, f) with average probabilistic guarantee $> p$ is signaling.*

Proof. Using the same arguments as in the proof of Theorem 2, we limit the analysis to AMOS and NEQ. For average case analysis, we consider these two games with input probability distribution $\pi_{xy} = \frac{1}{4}$ for every $(x, y) \in \{0, 1\}^2$. The success probability for Alice and Bob with this input distribution is then given by:

$$\Pr(\text{success}) = \frac{1}{4} \sum_{x,y} \sum_{a,b} \mathbb{1}_{\{\delta(a,b)=f(x,y)\}} \Pr(a, b|x, y)$$

First, we show that the protocol described in the proof of Theorem 2 for solving AMOS and NEQ has average success probability $\frac{3}{4}$. Indeed, the success probability of that protocol can be written as:

$$\Pr(\text{success}) = \frac{1}{4} \sum_{x,y} \Pr(\text{success}(x, y)) = \frac{1}{4}\Big(1 + \frac{2}{3} + \frac{2}{3} + \frac{2}{3}\Big) = \frac{3}{4}$$

because, the protocol always satisfies the first equation of both games, and satisfies each of the three other equations (of both games) with probability $\frac{2}{3}$.

Next, we show that a non-local box cannot solve AMOS or NEQ with average success probability greater than $\frac{3}{4}$. Indeed, we have

$$\begin{aligned} \Pr(\text{success}) &= \frac{1}{4}\Big[\Big(\sum_{(x,y)\neq(0,0)} \sum_{a,b} \mathbb{1}_{\{\delta(a,b)=f(x,y)\}} \Pr(a, b|x, y)\Big) \\ &+ \Big(\sum_{a,b} \mathbb{1}_{\{\delta(a,b)=f(0,0)\}} \Pr(a, b|0, 0)\Big)\Big] \end{aligned}$$

The first term is the same as the one analyzed in the proof of Theorem 2, where it was proved to be at most 2. The second term is at most $\sum_{ab} \Pr(a, b|0, 0) \leq 1$. Therefore, the average success probability for non-signaling boxes is at most $\frac{3}{4}$. □

The practical interest of the previous two theorems comes from their consequence to distributed quantum computing. By Corollary 1, we get:

Corollary 2. *Quantum correlations does not help for solving 2-player games that are not equivalent to any* XOR*-game. This limitation holds for both worst case, and average case analysis.*

6 Conclusion and open problems

6.1 Further work

One obvious generalization of the 2-player games is to consider games with more than two players, with IDs from 1 to $n \geq 2$. In the n-player game (δ, f), Player i receives boolean x_i as input, and must return a boolean a_i such that

$$\delta(a_1, \ldots, a_n) = f(x_1, \ldots, x_n)$$

in absence of communication between the players. As for two players, two classes of games deserve specific interest:

- XOR-games: $\delta(a_1, \ldots, a_n) = a_1 \oplus \cdots \oplus a_n$, for they generalize the CHSH game, and for they can be solved by a non-signaling box implementable by a circuit of PR boxes (see [7]);
- AND-games: $\delta(a_1, \ldots, a_n) = a_1 \wedge \cdots \wedge a_n$ for they correspond to the standard decision mechanism in the distributed computing literature (see, e.g., [22, 34]).

In particular, the n-player variant of AMOS is:

$$\bigwedge_{i=1}^{n} a_i = \bigwedge_{i \neq j} (\overline{x_i \wedge x_j}).$$

There exists a randomized protocol (see [22]), that is using individual random coins, and solves AMOS with success guarantee $\frac{\sqrt{5}-1}{2} \geq 0.61 > 1/2$. In this protocol, every selected player (i.e., with input 1) outputs 1 with probability p, and 0 with probability $1 - p$, where p is to be fixed later. Every non-selected player (i.e., with input 0) systematically outputs 0. Hence, if no players are selected, then the protocol always outputs the right answer. If one player is selected, then the protocol fails with probability $1 - p$, while if two or more players are selected then the protocol fails with probability at most p^2. Solving $p^2 = 1 - p$ results in picking the optimal probability $p^* = \frac{\sqrt{5}-1}{2}$.

On the other hand, we have seen in this paper that AMOS can be solved with success guarantee $\frac{2}{3} > p^*$ by two players applying a probabilistic protocol using shared randomness. One can actually

show that the same guarantee can be achieved with three players, by analyzing the following system

$$\begin{cases} a_0 \cdot b_0 \cdot c_0 = 1 \\ a_1 \cdot b_0 \cdot c_0 = 1 \\ a_0 \cdot b_1 \cdot c_0 = 1 \\ a_0 \cdot b_0 \cdot b_1 = 1 \\ a_1 \cdot b_1 \cdot c_0 = 0 \\ a_1 \cdot b_0 \cdot c_1 = 0 \\ a_0 \cdot b_1 \cdot c_1 = 0 \\ a_1 \cdot b_1 \cdot b_1 = 0 \end{cases}$$

which lists the eight equations for AMOS corresponding to the eight possible inputs of the games. Consider the protocol which solves that system after ignoring the second and seventh equations with probability $\frac{1}{3}$, the third and sixth with probability $\frac{1}{3}$, and the fourth and fifth with probability $\frac{1}{3}$. This protocol has success probability at least $\frac{2}{3}$ for every triple of inputs.

Unfortunately, the protocols for two and three players do not seem to extend easily to a higher number of players. For four players, we have designed an ad hoc probabilistic protocol using shared randomness, with success probability $\frac{9}{14} > \frac{\sqrt{5}-1}{2}$, but we failed to design a local protocol with success probability $\frac{2}{3}$. For more than four players, the ad hoc protocol could be generalized, but we have not identified a general pattern for it.

Instead, the lower bound $\frac{2}{3}$ on the probability of success for solving AMOS with non-signaling boxes established in this paper trivially extends to n players. We thus conclude by stating the following problem.

Open problem 1: Prove or disprove the existence of a shared-randomness probabilistic protocol that solves the n-player AMOS game with success probability $\frac{2}{3}$, for all $n \geq 2$.

6.2 Perspective

The results in this paper open new perspectives in term of distributed *checking*, a.k.a. distributed *verification*, which consists in having a set of, say, n processes deciding whether their global state (defined as the union of the local state of every individual process) satisfies some prescribed property, or not. The literature on this latter topic (see, e.g., [22, 23, 29, 34]) assumes a *decision* function δ which is applied to the set of individual decisions produced by the processes. Typically, each process should output a boolean b_i, and the global interpretation of the outputs is computed by

$$\delta(b_1, \ldots, b_n) = \bigwedge_{i=1}^{n} b_i \in \{\text{“yes”}, \text{“no”}\} \ .$$

The use of the AND operator is motivated by the requirement that the global state is valid if and only if all processes agree on some (local) validity condition. If this condition is locally violated somewhere in the system, then at least one process “raises an alarm” by outputting 0. However, recent advances in the theory of distributed checking [3, 24] demonstrate that using other decision functions δ significantly increases the power of the “checker”, or “verifier”. Our results show that some functions δ, in particular the classical AND operator, do not enable to use the power of quantum computing efficiently, compared to shared randomness, at least for 2-player games. In contrast, the exclusive-or operator is known to offer high potential, as far as distributed quantum

computing is concerned. In particular, [7] proved that every boolean function f on n independent players can be implemented by a circuit of PR boxes that output booleans b_i, $i = 1, \ldots, n$, satisfying

$$\bigoplus_{i=1}^{n} b_i = f(x_1, \ldots, x_n) \ .$$

The results in this paper give one more evidence of the impact of the decision function δ on the ability of "deciding" boolean predicates f.

6.3 Open problem

We have seen in Section 2.4 that quantum resources would not be of much help as far as 2-coloring the even-size ring C_{2n} distributedly is concerned. As pointed out in [27], the situation appears to be much more intriguing for 3-coloring the ring C_n. Indeed, the arguments used in [32] for establishing the $\Omega(\log^* n)$ lower bound do not seem to extend to non-signaling distributions. The main obstacle is that it appears to be difficult to bound the chromatic number of the configuration graph used in [32] whenever playing with non-signaling distributions. Hence, it is not clear whether or not there exists a quantum Las Vegas algorithm for 3-coloring C_n performing in $o(\log^* n)$ rounds. We believe that it is not the case. On the other hand, it may be the case that quantum resources would help in designing Monte Carlo algorithms with better success probabilities than classical (shared-randomness) algorithms.

Open problem 2: What is the complexity of 3-coloring C_n using quantum resources?

Acknowledgement. Both authors are thankful to Cyril Gavoille and Frédéric Magniez for fruitful discussions about the topic of this paper. The second author wants also to thank Zvi Lotker for early discussions on quantum distributed computing.

References

[1] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *J. Algorithms* **7**(4):567–583, 1986.

[2] H. Arfaoui and P. Fraigniaud. What can be computed without communications? In proc. *19th Int. Colloquium on Structural Information and Communication Complexity* (SIROCCO), pages 135-146, 2012.

[3] H. Arfaoui, P. Fraigniaud, and A. Pelc. Local Decision and Verification with Bounded-Size Outputs. In proc. *15th Int. Symp. on Stabilization, Safety, and Security of Distributed Systems* (SSS), 2013.

[4] L. Barenboim and M. Elkin. Distributed $(\Delta + 1)$-coloring in linear (in delta) time. In Proc. *41st ACM Symp. on Theory of computing (STOC)*, pages 111–120, 2009.

[5] L. Barenboim, M. Elkin, S. Pettie, and J. Schneider. The Locality of Distributed Symmetry Breaking. In Proc. *53rd IEEE Symp. on Foundations of Computer Science* (FOCS), pages 321–330, 2012.

[6] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. Nonlocal correlations as an information-theoretic resource. *Physical Review A* **71**(2):1-11, 2005.

[7] J. Barrett and S. Pironio. Popescu-Rohrlich correlations as a unit of nonlocality. *Phys. Rev. Lett.* **95**(14), 2005.

[8] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics* **1**(3):195–200, 1964.

[9] G. Brassard, A. Broadbent, and A. Tapp. Quantum pseudo-telepathy. *Foundations of Physics* **5**:18771907, 2005.

[10] A. Broadbent, A. Tapp. Can quantum mechanics help distributed computing? *SIGACT News* **39**(3): 67-76 (2008)

[11] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf. Non-locality and communication complexity. *Reviews of Modern Physics* **82**:665-698, 2010.

[12] H. Buhrman and H. Röhrig. Distributed Quantum Computing. In proc *28th International Symposium on Mathematical Foundations of Computer Science* (MFCS), LNCS 2747, pp. 120, 2003.

[13] B. S. Cirel'son. Quantum generalizations of bell's inequality. *Letters in Math. Phys.* **4**(2):93–100, 1980.

[14] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters* **23**(15):880–884, 1969.

[15] W. van Dam. Implausible Consequences of Superstrong Nonlocality. *Natural Computing* **12**(1): 9-12 (2013)

[16] V. Denchev and G. Pandurangan. Distributed quantum computing: a new frontier in distributed systems or science fiction? *SIGACT News* **39**(3): 77-95 (2008)

[17] B. Derbel, C. Gavoille, D. Peleg, and L. Viennot. On the locality of distributed sparse spanner construction. In proc. *27th ACM Symp. on Principles of Distributed Computing* (PODC), pages 273282, 2008.

[18] F. Dupuis, N. Gisin, A. Hasidim, A. Allan Méthot, and H. Pilpel. No nonlocal box is universal. *J. Math. Phys.* **48**(082107), 2007.

[19] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, **47**(10):777–780, 1935.

[20] M. Elkin. A near-optimal fully dynamic distributed algorithm for maintaining sparse spanners. In proc. *26th ACM Symp. on Principles of Distributed Computing* (PODC), pages 195204, 2007.

[21] M. Elkin, H. Klauck, D. Nanongkai, and G. Pandurangan. Quantum Lower Bounds for Distributed Network Computing. Tech. Report arXiv:1207.5211 (2013)

[22] P. Fraigniaud, A. Korman, and D. Peleg. Local distributed decision. In proc. *52nd Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pp 708-717, 2011.

[23] P. Fraigniaud, S. Rajsbaum, and C. Travers. Locality and checkability in wait-free computing. *25th International Symposium on Distributed Computing* (DISC), pp 333-347, 2011.

[24] P. Fraigniaud, S. Rajsbaum, and C. Travers. An Impossibility Result for Run-Time Monitoring. Submitted, 2013.

[25] S. J. Freedman and J. F. Clauser Experimental Test of Local Hidden-Variable Theories. *Phys. Rev. Lett.* **28**, 938941 (1972)

[26] C. Gavoille, R. Klasing, A. Kosowski, L. Kuszner, and A. Navarra. On the complexity of distributed graph coloring with local minimality constraints. *Networks* **54**(1): 12-19 (2009)

[27] C. Gavoille, A. Kosowski, and M. Markiewicz. What Can Be Observed Locally? In proc. *23rd Int. Symposium on Distributed Computing* (DISC), pages 243-257, 2009.

[28] Ghirardi, G. C. and Rimini, A. and Weber, T. A general argument against superluminal transmission through the quantum mechanical measurement process *Lett. Nuovo Cimento* **27**:293-298, 1980.

[29] A. Korman, S. Kutten, and D. Peleg. Proof labeling schemes. *Distributed Computing* **22**, (2010), 215–233.

[30] F. Kuhn. Weak graph colorings: distributed algorithms and applications. In Proc. *21st ACM Symp. on Parallel Algorithms and Architectures (SPAA)*, pages 138–144, 2009.

[31] F. Kuhn, T. Moscibroda, and R. Wattenhofer. What cannot be computed locally! In proc *23rd ACM Symp. on Principles of Distributed Computing* (PODC), pages 300-309, 2004.

[32] N. Linial. Locality in Distributed Graph Algorithms. *SIAM J. Comput.* **21**(1): 193-201 (1992)

[33] M. Luby. A simple parallel algorithm for the maximal independent set problem. *SIAM J. Comput.* **15**:1036–1053 (1986).

[34] M. Naor and L. Stockmeyer. What can be computed locally? *SIAM J. Comput.* **24**(6): 1259–1277 (1995).

[35] A. Panconesi and A. Srinivasan. On the Complexity of Distributed Network Decomposition. *J. Algorithms* **20**(2): 356-374 (1996).

[36] D. Peleg. Distributed computing: A locality-sensitive approach. SIAM, 2000.

[37] S. Popescu and D. Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics* **24**(3):379–385, 1994.

[38] J. Schneider and R. Wattenhofer. A new technique for distributed symmetry breaking. In Proc. *29th ACM Symp. on Principles of Distributed Computing* (PODC), 257-266, 2010.