Research Statement

Mika Göös
Institute for Advanced Study

1. Executive summary

I've been long fascinated by ***impossibility phenomena*** in mathematics and theoretical computer science: Gödel's incompleteness theorem, Turing's uncomputability of the halting problem, the $\mathsf{P} \neq \mathsf{NP}$ conjecture. My purpose in life is to contribute to this classical tradition. I want to discover new sources of impossibility by proving ***unconditional lower bounds*** on the amount of computational resources required to solve a given computational problem.

The main conceptual thrust in my research has been to develop general tools that allow one to **transfer** (or **"lift"**) lower-bound results from simple models of computation to analogous lower-bound results for much more powerful models. These new tools have allowed me and my coauthors to resolve several long-standing open problems in computational complexity theory.

The main motivating questions in my research are:

1. *Circuit complexity:* How much *time* is needed to solve a given problem?

2. *Communication complexity:* How many *bits of communication* are needed to solve a problem, when the input is distributed over several parties?

3. *Proof complexity:* Are there concisely stated theorems that admit no short proof?

2. Everything is communication

The most significant contributions in my PhD thesis (which won the 2017 EATCS Distinguished Dissertation Award) were in ***communication complexity***. What elevates this field beyond its superficial appearances (communication networks, distributed systems) are several surprising connections to other seemingly unrelated areas of theoretical computer science and mathematics. For example, some applications of my core communication results are as follows.

- **Circuit complexity.** In [8], we construct an n-variable monotone function in NP with near-maximal $\Omega(n/\log n)$ monotone circuit depth complexity (which is a measure of *parallel time* for monotone computations). This improves on the previous record of $\Omega(\sqrt{n})$ proved for the matching function by Raz and Wigderson (JACM 1992).

- **Proof complexity.** It has been known since the 1990s (Pudlák, Bonet–Pitassi–Raz) that one can extract monotone computations from many basic propositional proof systems, such as *Resolution* and *Cutting Planes*. In [1], we have shown that the *converse* holds. A corollary of this new connection is that *Gaussian elimination* does not admit an efficient monotone algorithm [6]. In particular, this result improves qualitatively on the monotone vs. non-monotone separation of Tardos (1988).

- **Combinatorial optimisation.** In [**3**], we construct an n-node graph whose independent set polytope requires linear programming formulations of size exponential in $\Omega(n/\log n)$. All previous lower bounds for explicit n-dimensional 0/1-polytopes (Fiorini et al., STOC 2012; Rothvoß, STOC 2014) were at best exponential in $\Theta(\sqrt{n})$.

- **Graph theory.** In [**2**], I disproved the *polynomial Alon–Saks–Seymour conjecture* that asked: Is the chromatic number of a graph always polynomial in the biclique partition number? (The conjecture holds for complete graphs by the basic Graham–Pollak theorem.)

A central theme in my research has been to simplify the analysis of communication protocols by exploiting new connections to **query complexity**, the study of *decision trees*. I have developed new query-to-communication "lifting" techniques that generically translate query complexity lower bounds into communication complexity lower bounds. My coauthors and I have used these new techniques to resolve several fundamental open problems in communication complexity, some dating back to Yao's original 1979 paper that started the field.

Other researchers, too, have found uses for these new techniques. For example, Ambainis et al. (JACM 2017) have extended our query-based methods to settle the famous Saks–Wigderson (FOCS 1986) conjecture, and Babichenko and Rubinstein (STOC 2017) have used our techniques to show that finding approximate Nash equilibria requires prohibitively many bits of communication. (Rubinstein and I later obtained near-optimal lower bounds for the latter problem [**12**].)

3.   My favourite result

In the paper [**2**] (which won the 2015 Machtey Award) I showed the first nontrivial lower bounds for the **Clique vs. Independent Set** (CIS) two-party communication game. This had been a long-standing open problem due to Yannakakis (STOC 1988). The CIS game, defined relative to an n-node graph $G = ([n], E)$, is simple to state: Alice holds a clique $x \subseteq [n]$ in G, Bob holds an independent set $y \subseteq [n]$ in G, and their goal is to decide whether x and y intersect. My result was to show that there exist graphs G such that any communication protocol (even co-nondeterministic) solving the CIS game on G must use $\omega(\log n)$ bits of communication. For comparison, Yannakakis showed that $O(\log^2 n)$ bits of communication always suffice.
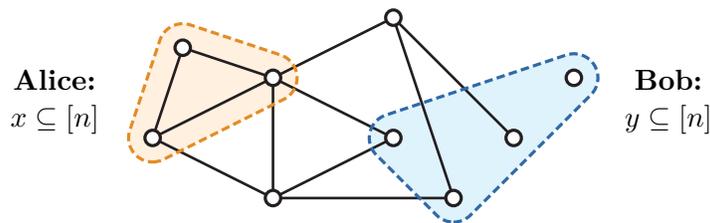
The reason why the CIS game is so interesting is because it admits many equivalent formulations. For example, my result is equivalent—via nontrivial reductions (Alon–Haviv, Bousquet et al.)—to the negation of the polynomial Alon–Saks–Seymour conjecture mentioned above. For another example, which was Yannakakis's original motivation, lower bounds for the CIS game imply lower bounds on the size of linear programming formulations.

In subsequent work [**9, 4**], we obtained a near-optimal lower bound $\tilde{\Omega}(\log^2 n)$ on the deterministic (and even randomised) communication complexity of the CIS game. In particular, this yields new lower bounds for the **log-rank conjecture** of Lovász and Saks (FOCS 1988), which is

Class	Query model	Communication model	Reference
P	deterministic	deterministic	Raz–McKenzie (1997)
NP	nondeterministic	nondeterministic	[7]
BPP	randomised	randomised	[10]
many	polynomial degree	rank	Sherstov, Shi–Zhu (2009)
P^{NP}	decision list	rectangle list	[5]
PLS	Resolution proof	monotone circuit	[1]
PPA	Nullstellensatz proof	monotone span program	Pitassi–Robere (2018)
	Sherali–Adams	LP extension complexity	Chen et al. (2014)
	sum-of-squares	SDP extension complexity	Lee et al. (2015)

Table 1: Known lifting theorems, classified according to domain of application.

perhaps the most famous open problem in communication complexity. Their conjecture posits that the deterministic communication complexity of any two-party function $F\colon \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ (Alice gets $x \in \mathcal{X}$, Bob gets $y \in \mathcal{Y}$, their goal is to compute $F(x,y)$) is polynomially related to the logarithm of the rank of F (viewed as a boolean matrix). We exhibit an F with deterministic communication complexity $\tilde{\Omega}(\log^2 \mathrm{rank}(F))$, improving the best previous bound due to Nisan, Kushilevitz, and Wigderson (FOCS 1994). All these lower bounds were proved using our aforementioned query-to-communication lifting techniques, which we next explain in more detail.

4. Query-to-communication lifting

Communication protocols are often tricky to analyse. A recurring theme in my research is to simplify this analysis by exploiting new connections between communication protocols and decision trees. To prove a communication lower bound, the basic idea is as follows.

− **Step 1:** Prove a general theorem stating that for a large class of communication problems F, any protocol for F can be efficiently *simulated* by a decision tree solving a related problem f.

− **Step 2:** Rule out efficient decision trees for f.

In other words, we can "lift" a query lower bound for f into a communication lower bound for F. This automates the task of proving communication lower bounds as we only need to show a problem-specific query lower bound for f (which is often relatively simple), and then invoke the general-purpose lifting theorem to obtain an analogous communication result.

The first fully developed lifting theorem was proved by Raz and McKenzie (FOCS 1997) for deterministic query/communication models. By now, there are many more such theorems; see Table 1. A remarkable recent example is the work of Lee, Raghavendra, and Steurer (STOC 2015), which remains the only known technique to prove lower bounds on the size of semi-definite programming formulations.

New lifting theorems. We have obtained the first lifting theorems for *randomised* [10], *nondeterministic* [7], *list-like* [5], and *dag-like* models [1]. In particular, before our work [10], the conspicuous lack of a randomised lifting theorem had been recognised as a major open problem, with some partial progress being published by several groups of researchers (ECCC 2017).

Further applications. In [9], we answered a basic question of Yao (STOC 1979), also posed in the Nisan–Kushilevitz (1997) textbook, concerning the relationship between communication complexity and the *partition number*. Perhaps the most basic observation in communication complexity is that a deterministic protocol of communication cost d that computes a function $F\colon \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ partitions the domain $\mathcal{X} \times \mathcal{Y}$ into at most 2^d *monochromatic rectangles*. A *rectangle* is a set $R = A \times B$ where $A \subseteq \mathcal{X}$, $B \subseteq \mathcal{Y}$, and we say R is *monochromatic* if F is constant on R. Hence, the logarithm of the *partition number* $\chi(F)$, defined as the least number of monochromatic rectangles needed to partition $\mathcal{X} \times \mathcal{Y}$, is a lower bound on deterministic communication complexity. We show that deterministic communication complexity of F is *not* characterised by $\Theta(\log \chi(F))$—there exists an F that requires $\tilde{\Omega}(\log^{1.5} \chi(F))$ bits of communication.

5. Future challenges

I believe the core motivating questions of complexity theory are *timeless*: they withstand the vicissitudes brought about by ever-changing trends in computer science. It is exciting to partake in developing a theory that is both fundamental and still relatively young (as fields of mathematics go). For future work, I mention two concrete example directions.

Structural communication complexity. Babai, Frankl, and Simon (FOCS 1986) introduced analogues of the standard complexity classes into communication complexity. While there have been many successes in separating complexity classes relative to oracles (via query complexity), the situation in communication complexity is much worse. For example, it is still open whether the polynomial hierarchy (PH) in communication complexity is strict. Razborov (1989) showed that understanding the communication PH is a necessary step in order to construct explicit ***rigid matrices*** in the sense of Valiant (1977)—an outstanding open problem in circuit complexity. Since Forster's (2002) result on the sign-rank (aka UPP communication complexity) of the inner-product function, no further progress has been made in proving lower bounds for increasingly stronger models of communication.

The frontier of our understanding lies with Arthur–Merlin (AM) communication protocols. We have written a survey [11] explaining the current state-of-the-art. Two recent papers (Bouland et al., FOCS 2017; Chattopadhyay–Mande, FOCS 2018) have already solved some of our open problems stated in the survey. In particular, the two papers pinpoint more precisely why the known lower-bound techniques fail for AM and related models.

Another broad future project is to develop a theory for the communication analogues of subclasses of TFNP (total NP search problems). We have written a survey about this in [6].

Circuit/proof complexity. Our recent work [1] imported lifting techniques to the study of *dag-like* models: monotone circuits and propositional proof systems (such as Resolution and Cutting Planes). If the long line of work on *tree-like* lifting theory is of any indication, there should be much to explore also in the dag-like setting. Can we prove that Cutting Planes cannot efficiently refute unsatisfiable linear systems over $\mathbb{F}_2$ (open since Cook–Coullard–Turán, 1987)? How about lower bounds for stronger proof systems, such as *Resolution over Cutting Planes* (open since Krajíček, 1998)?

References

[1] Ankit Garg, Mika Göös, Pritish Kamath, and Dmitrt Sokolov. Monotone circuit lower bounds from Resolution. In *Proceedings of the 50th Annual ACM Symposium on Theory of Computing* **(STOC'18)**, pages 902–911. ACM, 2018. Invited to Theory of Computing. doi:10.1145/3188745.3188838.

[2] Mika Göös. Lower bounds for clique vs. independent set. In *Proceedings of the 56th Symposium on Foundations of Computer Science* **(FOCS'15)**, pages 1066–1076. IEEE, 2015. Machtey award. doi:10.1109/FOCS.2015.69.

[3] Mika Göös, Rahul Jain, and Thomas Watson. Extension complexity of independent set polytopes. *SIAM Journal on Computing*, 47(1):241–269, 2018. Conference version in **(FOCS'16)**. doi:10.1137/16M109884X.

[4] Mika Göös, T.S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication vs. partition number. *ACM Transactions on Computation Theory*, 10(1):4:1–4:20, 2018. Conference version in **(ICALP'17)**. doi:10.1145/3170711.

[5] Mika Göös, Pritish Kamath, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for P^{NP}. In *Proceedings of the 32nd Computational Complexity Conference* **(CCC'17)**, pages 12:1–12:16. Schloss Dagstuhl, 2017. doi:10.4230/LIPIcs.CCC.2017.12.

[6] Mika Göös, Pritish Kamath, Robert Robere, and Dmitry Sokolov. Adventures in monotone complexity and TFNP. In *Proceedings of the 10th Innovations in Theoretical Computer Science* **(ITCS'19)**, 2019. To appear. URL: https://eccc.weizmann.ac.il/report/2018/163/.

[7] Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM Journal on Computing*, 45(5):1835–1869, 2016. Conference version in **(STOC'15)**. doi:10.1137/15M103145X.

[8] Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. *SIAM Journal on Computing*, pages 1778–1806, 2018. Conference version in **(STOC'14)**. doi:10.1137/16M1082007.

[9] Mika Göös, Toniann Pitassi, and Thomas Watson. Deterministic communication vs. partition number. In *Proceedings of the 56th Symposium on Foundations of Computer Science* **(FOCS'15)**, pages 1077–1088. IEEE, 2015. Invited to Special Issue. doi:10.1109/FOCS.2015.70.

[10] Mika Göös, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for BPP. In *Proceedings of the 58th Symposium on Foundations of Computer Science* **(FOCS'17)**, pages 132–143. IEEE, 2017. Invited to Special Issue. doi:10.1109/FOCS.2017.21.

[11] Mika Göös, Toniann Pitassi, and Thomas Watson. The landscape of communication complexity classes. *Computational Complexity*, 27(2):245–304, 2018. Conference version in **(ICALP'16)**. doi:10.1007/s00037-018-0166-6.

[12] Mika Göös and Aviad Rubinstein. Near-optimal communication lower bounds for approximate Nash equilibria. In *Proceedings of the 59th Symposium on Foundations of Computer Science* **(FOCS'18)**, pages 397–403. IEEE, 2018. Invited to Special Issue. arXiv:1805.06387.