Program

Tuesday, November 12, 2024

9:30 - 9:45 *Gathering, Registration, Coffee*

9:45 - 10:00 **Opening Remarks**

10:00 - 10:30 **Christoph Dürr -** *Randomized Binary and Tree Search under Pressure*

Abstract: We consider the problem of locating a hider in a path or a tree with edge queries. The twist is that we can do only k queries, which is too little to identify the precise location. Hence we study a game played between the searcher and the hider and ask questions such as how to compute the best response, what is the value of the game, what are the winning strategies.

This is joint work with Agustin Caracci José Verschae from the Ponitificia Universidad Católica de Chile.

10:30 - 11:00 **Yaniv Sadeh -** *Search Trees on Trees via Linear Programming*

Abstract: We consider the problem of search trees on trees (STTs) which generalizes binary search trees (BSTs) to searching nodes over a tree topology. In the BST case, the tree is a path. Golinsky's thesis (Tel Aviv University, 2023) proposes a linear-programming approach to the problem of computing a good static STT, shows that it computes an STT that is a 2-approximation to an optimal STT, and conjectures that it is in fact optimal. We study the LP approach further and show that the conjecture is false and that the LP approach is suboptimal for certain topologies. We also study additional aspects of this LP, among which is its dual form.

11:00 - 11:30 *Coffee break*

11:30 - 12:00 **Allan Borodin -** *Some new results for online interval selection*

Abstract: This talk will be in part a followup to my FILOFOCS 2023 talk. Yes, there are still things we don't know about online interval selection. In this talk I will discuss some recent results that I think raise interesting questions that go well beyond interval selection. In particular, we will discuss interval selection in the random order model when revoking (also called replacement) of intervals is allowed. This work will suggest a number of interesting directions for future research.

Joint work with Ben Cookson and Chris Karavasilis.

12:00 - 12:30 **Shahar Lewkowicz -** *List update with prediction*

Abstract: List Update is a fundamental problem in online algorithms, with a well-known 2-competitive algorithm that moves every requested element to the front. Randomization can slightly improve the competitive ratio to 1.6, but not beyond 1.5. However, practical inputs are not adversarial and one hopes to do better, particularly when additional information from a machine learning oracle is available. With access to predictions, the goal is to incur only a slight overhead compared to the prediction's accuracy, avoiding significant costs in case of substantial deviation.

We propose a $(1+\epsilon)$-smooth randomized algorithm, offering robustness of $O(1/\epsilon^4)$. This guarantees that the algorithm never exceeds a cost greater than $1+\epsilon$ times the prediction

cost, while maintaining a bound within $O(1/\epsilon^4)$ of the optimal cost for every possible sequence. In cases where no paid swaps are permitted for the prediction, we can improve robustness to $O(1/\epsilon^2)$ while retaining $1+\epsilon$ smoothness. We complement these findings by demonstrating a lower bound of $\Omega(1/\epsilon)$ on the robustness for deterministic algorithms and $\Omega(\log(1/\epsilon))$ for randomized ones. Finally, the experiments we have made show that our algorithms perform better than the standard competitive algorithms for this problem.

Joint Work with Yossi Azar and Varun Suriyanarayana.

12:30 - 14:00 *Lunch* (On your own)

14:00 - 14:30 **Sergio Rasjbaum -** *Distributed Computability and Continuous Maps*

Abstract: The celebrated Herlihy and Shavit Asynchronous Computability theorem characterizes the tasks that are wait-free solvable in a shared memory read/write system in terms of certain subdivisions relating the input and output complex of a task. While this result is fundamental, it is not as informative as the corresponding result for colorless tasks, which directly connects wait-free solvability and topology: it states that a colorless task is solvable if and only if there is a continuous map from the input to the output complex respecting the input/output relation of the task. A self-contained, intuitive introduction to this topic will be presented, explaining these theorems, some of their important consequences, and the efforts that have been done at presenting general task solvability characterizations in terms of continuous maps.

14:30 - 15:00 **Yuval Emek -** *On the power of Graphical reconfigurable circuits*

Abstract: In this talk, we introduce the graphical reconfigurable circuits (GRC) model as an abstraction for distributed graph algorithms whose communication scheme is based on local mechanisms that collectively construct long-range reconfigurable channels (this is an extension to general graphs of a distributed computational model recently introduced by Feldmann et al. (JCB 2022) for hexagonal grids). The crux of the GRC model lies in its modest assumptions: (1) the individual nodes are computationally weak, with state space bounded independently of any global graph parameter; and (2) the reconfigurable communication channels are highly restrictive, only carrying information-less signals (a.k.a. beeps). Despite these modest assumptions, we prove that GRC algorithms can solve many important distributed tasks efficiently, i.e., in polylogarithmic time. On the negative side, we establish various runtime lower bounds, proving that for other tasks, GRC algorithms (if they exist) are doomed to be slow.

The talk is based on a recent joint work with Yuval Gil and Noga Harlev (DISC 2024), which belongs to a more general research agenda dedicated to uniform distributed computational models.

15:00 - 15:30 **Dror Rawitz -** *On Key Parameters Affecting the Realizability of Degree Sequences*

Abstract: The Degree Realization problem with respect to a graph family $\mathcal{F}$ is defined as follows. The input is a sequence d of n positive integers, and the goal is to decide whether there exists a graph $G \in \mathcal{F}$ whose degrees correspond to d. The main challenges are to provide a characterization of all the sequences that admit a realization in $\mathcal{F}$ and to design an efficient algorithm that constructs one of the possible realizations, if one exists.

We discuss the parameters affecting the realizability of degree sequences by restricted classes of sparse graph, including planar graphs, outer-planar graphs, and some of their subclasses (e.g., 2-trees and cactus graphs). Our motivation is that the problems of characterizing *planaric and* outer-planaric degree sequences are open. For planar graphs, for example, an

obvious consequence of Euler's theorem is that $\sum_i d_i \leq 6n - 12$ is a *necessary condition for d to be planaric. This condition can be improved to $\sum_i d_i \leq 6n - 12 - 2\omega_2 - 4\omega_1$, where ω_i is the multiplicity of i in d. In this talk we present* sufficient conditions for planaric and outer-planaric sequences. For example, we show that a graphic sequence d is planaric if $\sum d \leq 4n - 4 - 2\omega_1$.

Joint work with Amotz Bar-Noy, Toni Böhnlein, David Peleg, and Yingli Ran.

15:30 - 16:00 *Coffee break*

16:00 - 16:30 **Jonas Ellert -** *Faster Two-Dimensional Pattern Matching with k Mismatches*

Abstract: The classical pattern matching asks for locating all occurrences of one string, called the pattern, in another, called the text, where a string is simply a sequence of characters. Due to the potential practical applications, it is desirable to seek approximate occurrences, for example by bounding the number of mismatches. This problem has been extensively studied, and by now we have a good understanding of the best possible time complexity as a function of n (length of the text), m (length of the pattern), and k (number of mismatches). In particular, we know that for $k = O(\sqrt{m})$, we can achieve quasi-linear time complexity [Gawrychowski and Uznanski, ICALP 2018].

We consider a natural generalisation of the approximate pattern matching problem to two-dimensional strings, which are simply square arrays of characters. In the approximate two-dimensional pattern matching, we are given a pattern of size $m \times m$ and a text of size $n \times n$, and ask for all locations in the text where the pattern matches with at most k mismatches. The asymptotically fastest algorithm for this algorithm works in $O(kn^2)$ time [Amir and Landau, TCS 1991]. We provide a new insight into two-dimensional periodicity to improve on these 30-years old bounds. Our algorithm works in $\tilde{O}((m^2 + mk^{5/4})n^2/m^2)$ time, which is $\tilde{O}(n^2)$ for $k = O(m^{4/5})$.

16:30 - 17:00 **Zvi Lotker -** *Computational History: New Paths in Digital Humanities*

Abstract: In this talk, we will explore how computational methods and mathematical models can illuminate historical patterns, turning the complexities of human history into quantifiable insights. From network theory applied to social connections in ancient societies to probabilistic models that reconstruct missing links in the historical record, we will examine how mathematics allows us to decipher the structure and evolution of historical events. We will also discuss how to express historical patterns in mathematical language, providing a framework for rigorous computational analysis. The material for this talk is based on my book, *Computational History*, which explores these themes in depth and highlights the potential of mathematics to enhance historical inquiry.

Wednesday, November 13, 2024

9:00 - 9:30 *Gathering, Coffee*

9:30 - 10:00 **Srindhi Nagendra -** *Reward Augmentation in Reinforcement Learning for Testing Distributed Systems*

Abstract: Bugs in popular distributed protocol implementations have been the source of many downtimes in popular internet services. We describe a randomized testing approach for distributed protocol implementations based on reinforcement learning. Since the natural reward structure is very sparse, the key to successful exploration in reinforcement learning is reward augmentation. We show two different techniques that build on one another. First, we provide a decaying exploration bonus based on the discovery of new states – the reward decays as the same state is visited multiple times. The exploration bonus captures the intuition from coverage-guided fuzzing of prioritizing new coverage points; in contrast to other schemes, we show that taking the maximum of the bonus and the Q-value leads to more effective exploration. Second, we provide waypoints to the algorithm as a sequence of predicates that capture interesting semantic scenarios. Waypoints exploit designer insight about the protocol and guide the exploration to "nteresting" parts of the state space. Our reward structure ensures that new episodes can reliably get to deep interesting states even without execution caching. We have implemented our algorithm in Go. Our evaluation on three large benchmarks (RedisRaft, Etcd, and RSL) shows that our algorithm can significantly outperform baseline approaches in terms of coverage and bug finding.

10:00 - 10:30 **Junyao Zhao -** *Strategizing against No-Regret Learners in First-Price Auctions*

Abstract: We study repeated first-price auctions and general repeated Bayesian games between two players, where one player, the learner, employs a no-regret learning algorithm, and the other player, the optimizer, knowing the learner's algorithm, strategizes to maximize its own utility. For a commonly used class of no-regret learning algorithms called mean-based algorithms, we show that (i) in full-information first-price auctions, the optimizer cannot get more than the Stackelberg utility – a standard benchmark in the literature, but (ii) in Bayesian first-price auctions, there are instances where the optimizer can achieve much higher than the Stackelberg utility.

On the other hand, Mansour et al. (2022) showed that a more sophisticated class of algorithms called no-polytope-swap-regret algorithms are sufficient to cap the optimizer's utility at the Stackelberg utility in any repeated Bayesian game (including Bayesian first-price auctions), and they pose the open question whether no-polytope-swap-regret algorithms are necessary to cap the optimizer's utility. For general Bayesian games, under a reasonable and necessary condition, we prove that no-polytope-swap-regret algorithms are indeed necessary to cap the optimizer's utility.

10:30 - 11:00 *Coffee break*

11:00 - 11:30 **Yoav Gal Tzur -** *Combinatorial multi-action contracts*

Abstract: Algorithmic contract design is an emerging frontier in the intersection of economics and computation, with combinatorial contracts being a core problem in this domain. A central model within combinatorial contracts explores a setting where a principal delegates the execution of a project, which can either succeed or fail, to an agent. The agent can choose any subset among a given set of costly actions, where every subset is associated with a success probability of the project, given by a set function f. The principal incentivizes the agent through a contract that specifies the payment upon success of the project.

We provide results on the complexity of the associated optimization problem — finding the contract that maximizes the principal's utility. On the positive side, we show that when the principal has access to a demand oracle for f, the optimal contract can be computed efficiently, if there are poly-many breakpoints in the agent's piece-wise utility function. A direct corollary is a poly-time algorithm for the optimal contract in settings where f is supermodular. On the negative side, we show that for submodular f, exponentially-many demand oracle calls may be required to find the optimal contract.

Joint work with Paul Dütting, Michal Feldman and Aviad Rubinstein.

11:30 - 12:00 **Tomek Ponitka -** *Pseudo-Dimension of Contracts*

Abstract: Algorithmic contract design is an emerging frontier at the intersection of economics and computation, studying scenarios where a principal delegates the execution of a costly project to an agent. Most research has focused on computing the optimal contract in situations where the principal has full distributional knowledge of the contract setting — a highly unrealistic assumption in many real-world situations. In this work we relax this assumption, and take a learning approach to the contract design problem in settings where the underlying distribution is unknown. We devise learning algorithms that provide nearly-optimal contracts based on samples from the underlying distribution, for three major classes of contracts. Similar to a learning approach to mechanism design, our analysis is based on tight bounds that we establish on the pseudo-dimension of contracts.

12:00 - 12:30 **Michal Feldman -** *Ambiguous contracts*

Abstract: In this work we explore the deliberate infusion of ambiguity into the design of contracts. We show that when the agent is ambiguity-averse and chooses an action that maximizes their max-min utility, then the principal can strictly gain from using an ambiguous contract. We provide insights into the structure of optimal contracts, and establish that optimal ambiguous contracts are composed of simple contracts. We also provide a characterization of ambiguity-proof classes of contracts. Finally, we show that when the agent considers mixed strategies, then there is no advantage in using an ambiguous contract.

12:30 - 14:00 *Lunch* (On your own)

14:00 - 14:30 **Geoffroy Couteau -** *Fast Public-Key Silent OT and More from Constrained Naor-Reingold*

Abstract: Pseudorandom Correlation Functions (PCFs) allow two parties, given correlated evaluation keys, to locally generate arbitrarily many pseudorandom correlated strings, e.g. Oblivious Transfer (OT) correlations, which can then be used by the two parties to jointly run secure computation protocols.

In this work, we provide a novel and simple approach for constructing PCFs for OT correlation, by relying on constrained pseudorandom functions for a class of constraints containing a weak pseudorandom function (wPRF). We then show that tweaking the Naor-Reingold pseudorandom function and relying on low-complexity pseudorandom functions allow us to instantiate our paradigm. We further extend our ideas to obtain efficient public-key PCFs, which allow the distribution of correlated keys between parties to be non-interactive: each party can generate a pair of public/secret keys, and any pair of parties can locally derive their correlated evaluation key by combining their secret key with the other party's public key.

In addition to these theoretical contributions, we detail various optimizations and provide concrete instantiations of our paradigm relying on the Boneh-Ishai-Passelègue-Sahai-Wu wPRF and the Goldreich-Applebaum-Raykov wPRF. Putting everything together, we obtain public-key PCFs with a throughput of 15k-40k OT/s, which is of a similar order of magnitude to the

state-of-the-art interactive PCFs and about 4 orders of magnitude faster than state-of-the-art public-key PCFs.

14:30 - 15:00 **Christina Boura -** *Alternative Key Schedules for the AES*

Abstract: The AES block cipher is today the most important and analyzed symmetric cryptographic algorithm. While all versions of the AES are known to be secure in the single-key setting, this is not the case in the related-key scenario. In this talk we try to answer the question whether the AES would resist better differential-like related-key attacks if the key schedule was different. For this, we search for alternative permutation-based key schedules by extending previous works. We also develop different approaches together with MILP-based tools to find good permutations that could be used as the key schedule for AES-128, AES-192 and AES-256. Our methods permit to find permutations that outperform previous results for AES-128. Moreover, our new approach based on two MILP models that call one another allow us to handle a larger search space and thus to search for alternative key schedules for the two bigger versions of AES.

15:00 - 15:30 **Eden Aldema Tshuva -** *Fully Local Distributed Arguments*

Abstract: Distributed certification is a proof system for detecting illegal network states or improper execution of distributed algorithms. A certification scheme consists of a proving algorithm, which assigns a certificate to each node, and a verification algorithm where nodes use these certificates to decide whether to accept or reject. The system must ensure that all nodes accept if and only if the network is in a legal state, adhering to the principles of completeness and soundness. The main goal is to design a scheme where the verification process is local and the certificates are succinct, while using as efficient as possible proving algorithm. In cryptographic proof systems, the soundness requirement is often relaxed to computational soundness, where soundness is guaranteed only against computationally bounded adversaries. Computationally sound proof systems are called arguments. Recently, Aldema Tshuva, Boyle, Cohen, Moran, and Oshman (TCC 2023) showed that succinct distributed arguments can be used to enable any polynomially bounded distributed algorithm to certify its execution with polylogarithmic-length certificates. However, their approach required a global communication phase, adding $O(D)$ communication rounds in networks of diameter D, which limits its applicability to local algorithms.

In this work, we give the first construction of a fully local succinct distributed argument system, where the prover and the verifier are both local. We show that a distributed algorithm that runs in R rounds, has polynomial local computation, and messages of B bits each can be compiled into a self-certifying algorithm that runs in $R + \text{polylog(n)}$ rounds and sends messages of size $B + \text{polylog(n)}$, with certificates of length $\text{polylog}(n)$. This construction has several applications, including self-certification for local algorithms, ongoing certification of long-lived algorithms, and efficient local mending of the certificates when the network changes.

15:30 - 16:00 *Coffee Break*

16:00 - 16:30 **Alex B. Grilo -** *The role of piracy in quantum proofs*

Abstract: A well-known feature of quantum information is that it cannot, in general, be cloned. Recently, a number of quantum-enabled information-processing tasks have demonstrated various forms of uncloneability; among these forms, piracy is an adversarial model that gives maximal power to the adversary, in controlling both a cloning-type attack, as well as the evaluation/verification stage. Here, we initiate the study of anti-piracy proof systems, which are proof systems that inherently prevent piracy attacks. We define anti-piracy proof systems, demonstrate such a proof system for an oracle problem, and also describe a candidate

anti-piracy proof system for NP. We also study quantum proof systems that are cloneable and settle the famous QMA vs. QMA(2) debate in this setting. Lastly, we discuss how one can approach the QMA vs. QCMA question, by studying its cloneable variants.

This is a joint work with Anne Broadbent, Supartha Podder and Jamie Sikora.

16:30 - 17:00 **Pouria Fallahpour -** *Quantum Oblivious LWE Sampling and Insecurity of Standard Model Lattice-Based SNARKs*

Abstract: The Learning With Errors (LWE) problem asks to find a vector s from an input of the form $(A, b = As + e)$, for a matrix A and a vector e that has small-magnitude entries. In this talk, I focus on the task of sampling LWE instances. As these are extremely sparse in their range, it may seem plausible that the only way to proceed is to first create s and e and then set $b = As+e$. In particular, such an instance sampler knows the solution. This raises the question whether it is possible to obliviously sample $(A, As+e)$, namely, without knowing the underlying secret s. A variant of the assumption that oblivious LWE sampling is hard has been used in a series of works to analyze the security of candidate constructions of Succinct Non-interactive Arguments of Knowledge (SNARKs). As the assumption is related to LWE, these SNARKs have been conjectured to be secure in the presence of quantum adversaries. The main focus of the talk is a quantum polynomial-time algorithm that samples well-distributed LWE instances while provably not knowing the solution, under the assumption that LWE is hard. Moreover, the approach works for a vast range of LWE parameterizations, including those used in the above-mentioned SNARKs. This invalidates the assumptions used in their security analyses, although it does not yield attacks against the constructions themselves.

17:00 - 17:30 **Philip Verduyn Lunel -** *Relating non-local quantum computation to information theoretic cryptography*

Abstract: Non-local quantum computation (NLQC) is a cheating strategy for position-verification schemes. In this talk, we connect NLQC to the wider context of information theoretic cryptography by relating it to a number of other cryptographic primitives. We show one special case of NLQC, known as f-routing, is equivalent to the quantum analogue of the conditional disclosure of secrets (CDS) primitive, where by equivalent we mean that a protocol for one task gives a protocol for the other with only small overhead in resource cost. We further consider another special case of NLQC, which we call coherent function evaluation (CFE), and show CFE protocols induce similarly efficient protocols for the private simultaneous message passing (PSM) scenario. By relating position-verification to these cryptographic primitives, a number of results in the information theoretic cryptography literature give new implications for NLQC, and vice versa. For NLQC, we obtain the first sub-exponential upper bounds on the worst case cost of f-routing, and the first example of an efficient f-routing strategy for a problem believed to be outside P/poly, linear lower bounds on entanglement for CDS in the quantum setting, linear lower bounds on communication cost of CFE, and efficient protocols for CDS in the quantum setting for functions that can be computed with quantum circuits of low T depth.

18:30 - *Social Event* (to be confirmed)

Thursday, November 14, 2024

— — — Change of Building — — —

9:00 - 9:30 *Gathering, Coffee*

9:30 - 10:00 **Robert Krauthgamer -** *Cut Sparsification and Succinct Representation of Submodular Hypergraphs*

Abstract: Cut sparsification of hypergraphs is a widely applied method, where all cuts of a hypergraph H are approximated within $1+/-\epsilon$ factor by a hypergraph H' with few hyperedges. It was recently generalized to a setting where the cost of cutting each hyperedge $e$is provided by a given function g_e, that is usually assumed to be submodular.

I will present new bounds for sparsification of submodular hypergraphs, including the first polynomial bound for all submodular functions, and further improvements for various families of submodular functions. I will focus on the common setting, where the sparsifier H? is a reweighted sub-hypergraph of H, and its size is measured by the number of hyperedges, but will address also a more general notion of representing H' succinctly, where size is measured in bits.

Joint work with Yotam Kenneth.

10:00 - 10:30 **Alin Bostan -** *A Fast Algorithm for Computing Terms in Linearly Recurring Sequences and Its Application to Quasi-Optimal Composition*

Abstract: In algebraic complexity theory, it has been known since the 1970s that most operations (multiplication, division, interpolation, extended gcd, resultant, etc.) on univariate polynomials can be performed in quasi-linear time with respect to the degree. The two important ingredients for achieving such quasi-optimal complexity bounds are Fast Fourier Transform (FFT) techniques and the use of the divide-and-conquer paradigm.

Another basic operation is composition: given two polynomials f and g of degree less than n, compute $f(g(x))\mathrm{mod}x^n$. For this operation, no quasi-linear time algorithm was known before 2024. I will present a breakthrough result due to Yasunori Kinoshita and Baitian Li (FOCS 2024) that solves this long-standing open problem. Their algorithm builds on a recent algorithm due to Ryuhei Mori and myself (SOSA 2021) that efficiently solves a quite different problem, namely computing a selected coefficient in the Taylor expansion of a rational function $P(x)/Q(x)$, where P and Q two polynomials of degree less than n.

10:30 - 11:00 *Coffee break*

11:00 - 11:30 **Carola Doerr -** *Generating Point Sets of Small Star Discrepancy*

Abstract: In this talk, we will present different approaches to construct low discrepancy point sets. We first introduce mathematical programming formulations to construct point sets with optimal L_∞ star discrepancy values in dimension 2 for up to 21 points and up to 8 points in dimension 3. We show that these optimal sets have a far lower discrepancy than the previous references. More importantly, they present a very different structure.

We will then discuss extensions of this approach to obtain good, but not provably optimal, point sets and show that there is much room for improvement over state-of- he-art constructions. We also show that additional symmetry requirements can be satisfied at very small loss in discrepancy value. Finally, we will discuss some recent advances using graph neural networks by Rusch et al. [RKBR24] (PNAS) and how we outperform these constructions using classic optimization approaches [CDKP24].

The presentation is based on joint work with François Clément (University of Washington, US), Kathrin Klamroth (University of Wupperal, Germany), and Luis Paquete (Universirty of Coimbra, Portugal).

11:30 - 12:00 **Omri Weinstein -** *A multiplicative-Weights Method for Sparse Recovery*

Abstract: We develop a new approach to sparse minimization, which mitigates the limitation of classical greedy algorithms, most notably Orthogonal Matching Pursuit (OMP, Natarajan'95). Our approach is inspired by an elegant combinatorial algorithm of Clarkson ('95) for finding a subset of tight constraints for "all" linear programs (#constraints >> #vars). We generalize this randomized algorithm to the (non-convex) setting of sparse recovery, and show that in practice it outperforms state-of-art greedy algorithms (OMP, IHT).

Joint work with Erez Badash (HUJI).

12:00 - 12:30 **Adrian Vladu -** *Sparse Recovery Methods for Training Sparse Deep Learning Models*

Abstract: Sparse recovery has long been a key area of study in the context of compressed sensing, where the goal is to reconstruct a sparse signal from limited linear measurements. This problem is equivalent to identifying sparse solutions to underdetermined linear systems. Here, we extend these principles to a broader class of convex optimization problems, focusing on scenarios where the feasible set may itself be non-convex, such as sparse or quantized domains. This generalized framework not only provides a powerful, principled approach for tackling optimization over non-convex sets but also shows remarkable practical utility, especially in deep learning.

In the context of deep learning, sparsity is often pursued to create efficient models that retain high accuracy while requiring significantly fewer resources for inference. Surprisingly, this approach allows for aggressive sparsification, removing up to 95% of model weights with negligible impact on accuracy.

I will discuss recent work that systematically explores the landscape of sparse optimization, presenting a new accelerated method that achieves faster convergence while producing solutions with improved sparsity relative to standard (projected) gradient descent. These advancements offer a promising pathway to scalable, sparse optimization in machine learning.

12:30 - 14:00 *Lunch* (On your own)

14:00 - 14:30 **Claire Mathieu -** *Testing the similarity between two distributions in streaming*

Abstract: TBA

14:30 - 15:00 **Alexander Shen -** *From information complexity to aggregating opinions*

Abstract: The inequality saying that internal information complexity for a communication protocol does not exceed the external one can be rephrased as $I(\pi) \leq I(\pi|\alpha) + I(\pi|\beta)$, where π is a sequence of messages produced by the protocol on inputs α and β. Romashchenko and Zimand proved a Kolmogorov complexity version of this statement; it says that for an arbitrary combinatorial rectangle Π in $\mathbb{B}^n \times \mathbb{B}^n$ and for every pair $(x, y) in \Pi$ we have $C(\Pi) \leq C(\Pi||x) + C(\Pi||y) + O(i(Pi) + \log n)$, where $i(\Pi)$ is the maximal complexity of Π given its elements. This inequality has combinatorial translation that can be used (as Kozachinskiy and Steifer found) to prove the following result about opinion aggregation.

Imagine a group of people traveling in a train. At every stop the train conductor comes and asks passengers whether they want the heating to be on or off (till the next stop). Some people want it on, some other want it off, some do not care and are happy with both options. Every decision of the conductor makes some passengers unhappy (those who wanted the opposite option), and this is unavoidable if there are conflicting requests, but the conductor wants to

minimize the *maximal unhappiness* among passengers. (The unhappiness of a passenger is the number of her requests that were not fulfilled.) This can be difficult: for example, if two passengers have conflicting requests all the time (on every stop), then for t stops at least one of them has unhappiness at least $t/2$ (for obvious reasons). However, the following result is true for some polynomial p:

If every pair of passengers has a conflict at most once, then the conductor can guarantee that the maximal unhappiness after t stops does not exceed $\sqrt{t} \cdot p(\log n)$ where n is a number of passengers. (This is an existence result, the best bound known for explicit algorithms has $t^{2/3}$ instead of $\sqrt{t}$.)

15:00 - 15:30 **Gregory Kucherov -** *Better Space-Time-Robustness Trade-Offs for Set Reconciliation*

Abstract: We consider the problem of reconstructing the symmetric difference between similar sets from their representations (sketches) of size linear in the number of differences. Exact solutions to this problem are based on error-correcting coding techniques and suffer from a large decoding time. Existing probabilistic solutions based on Invertible Bloom Lookup Tables (IBLTs) are time-efficient but offer insufficient success guarantees for many applications. Here we propose a tunable trade-off between the two approaches combining the efficiency of IBLTs with exponentially decreasing failure probability. The proof relies on a refined analysis of IBLTs proposed in (Bæk Tejs Houen et al. SOSA 2023) which has an independent interest. We also propose a modification of our algorithm that enables telling apart the elements of each set in the symmetric difference.

joint work with Djamal Belazzougui (CERIST Algeria) and Stefan Walzer (KIT Germany).

15:30 - 16:00 *Coffee Break*

16:00 - 16:30 **Dani Dorfman -** *Faster all-pairs optimal car routing*

Abstract: We present a randomized $\tilde{O}(n^{3.5})$-time algorithm for computing *optimal energetic paths* for an electric car between all pairs of vertices in an n-vertex directed graph with positive and negative *costs*, or *gains*, which are defined to be the negatives of the costs. The optimal energetic paths are finite and well-defined even if the graph contains negative-cost, or equivalently, positive-gain, cycles. This makes the problem much more challenging than standard shortest paths problems.

More specifically, for every two vertices s and t in the graph, the algorithm computes $\alpha_B(s,t)$, the maximum amount of charge the car can reach t with, if it starts at s with full battery, i.e., with charge B, where B is the capacity of the battery. The algorithm also outputs a concise description of the optimal energetic paths that achieve these values. In the presence of positive-gain cycles, optimal paths are not necessarily simple. This improves on a previous $\tilde{O}(mn^2)$-time algorithm of Dorfman et al. [ESA 2024] for the problem.

The *gain* of an arc is the amount of charge added to the battery of the car when traversing the arc. The charge in the battery can never exceed the capacity B of the battery and can never be negative. An arc of positive gain may correspond, for example, to a downhill road segment, while an arc with a negative gain may correspond to an uphill segment. A positive-gain cycle, if one exists, can be used in certain cases to charge the battery to its capacity. This makes the problem more interesting and more challenging. As mentioned, optimal energetic paths are well-defined even in the presence of positive-gain cycles. Positive-gain cycles may arise when certain road segments have magnetic charging strips, or when the electric car has solar panels.

Combined with a result of Dorfman et al. [SOSA 2024], this also provides a randomized $\tilde{O}(n^{3.5})$-time algorithm for computing *minimum-cost paths* between all pairs of vertices in an

n-vertex graph when the battery can be externally recharged, at varying costs, at intermediate vertices.

Joint work with Haim Kaplan, Mikkel Thorup, Uri Zwick and Robert E. Tarjan .

16:30 - 17:00 **Yossi Azar -** *Online predictions and* 1*-smoothness*

ABSTRACT: We consider learning augmented problems, where our goal is to be at least as good as an algorithm that "follows the prediction". Our prediction at each step is the action taken by an optimal algorithm (or any information that yields such an action). An algorithm is called 1-smooth if its cost (value) is as low (high) as the algorithm that follows the prediction. For the packet scheduling with deadlines problem (where the goal is to maximize the total value of transmitted packets), we provide 1-smooth and 3-robust algorithm. That means an algorithm that always achieves a value which is as large as "following the prediction" but never worse than 3 times the optimal solution for every possible sequence and prediction.

Joint work with Or Vardi.

Friday, November 15, 2024

Special Machine Learning Day
organized in cooperation with ERC project COLT-MDP

8:45 - 9:00 *Gathering and Greetings, Coffee*

9:00 - 9:30 **Francis Bach -** *An alternative view of denoising diffusion models*

Abstract: Denoising diffusion models have led to impressive generative models in many domains. In this talk, I will present recent progress, with a focus on formulations that do not involve stochastic differential equations.

9:30 - 10:00 **Julia Kempe -** *Synthetic Data - Friend or Foe*

Abstract: As AI model size grows, neural *scaling laws* have become a crucial tool to predict the improvements of large models when increasing capacity and the size of original (human or natural) training data. Yet, the widespread use of popular models means that the ecosystem of online data and text will co-evolve to progressively contain increased amounts of synthesized data. In this talk we ask: *How will the scaling laws change in the inevitable regime where synthetic data makes its way into the training corpus?* Will future models still improve, or be doomed to degenerate up to total *(model) collapse*? We develop a theoretical framework of model collapse through the lens of scaling laws. We discover a wide range of decay phenomena, analyzing loss of scaling, shifted scaling with number of generations, the "un-learning" of skills, and grokking when mixing human and synthesized data. Our theory is validated by large-scale experiments with a transformer on an arithmetic task and text generation using the LLM Llama2. We also propose solutions to circumvent degradation in learning by pruning the generated data.

10:00 - 10:30 **Shay Moran -** *On Differentially Private Linear Algebra*

Abstract: We introduce efficient differentially private (DP) algorithms for several linear algebraic tasks, including solving linear equalities over arbitrary fields, linear inequalities over the reals, and computing affine spans and convex hulls. As an application, we obtain efficient DP algorithms for learning halfspaces and affine subspaces. Our algorithms addressing equalities satisfy strong polynomial-time efficiency, whereas those addressing inequalities achieve only weak polynomial-time efficiency. Furthermore, this distinction is inevitable: no DP algorithm for linear programming can be strongly polynomial-time efficient.

10:30 - 11:00 *Coffee break*

11:00 - 11:30 **Roi Livni -** *The information complexity of learning in stochastic convex optimization*

Abstract: An key principle in learning is to learn the distribution, not the data, and avoid overfitting. This principle has been formalized through various generalization bounds such as compression schemes, PAC Bayes bounds, and other information-theoretic generalization bounds. These bounds demonstrate how models that avoid memorizing the whole dataset can successfully generalize to unseen examples.

In this work we revisit this paradigm and demonstrate that in overparameterized setups (where number of parameters is far greater than number of examples) the relation between generalization and memorization might be more complex. We focus on the setup of stochastic convex optimization, and we show that successful generalization often requires heavy memorization of the data, which makes standard information-theoretic and compression-type generalization bounds vacuous even in setups where learning is tractable.

Based on joint work with: Idan Attias, Gintare Karolina Dziugaite, Mahdi Haghifam, and Daniel Roy.

11:30 - 12:00 **Ohad Shamir -** *Overfitting: The good, the bad and the ugly*

Abstract: In classical machine learning theory, overfitting the training data is invariably associated with a failure to learn successfully. However, in recent years it has become increasingly recognized that the situation is far more nuanced, and that overfitting can lead to a spectrum of learning outcomes, ranging from perfectly successful learning to catastrophic failure, as well as interesting intermediate cases. In this talk, I'll describe several recent results and examples of this behavior, in the context of neural networks, linear and kernel predictors.

Based on joint works with Daniel Barzilai, Guy Kornowski. and Gilad Yehudai.

12:00 - 12:30 **Nadav Cohen -** *Implicit Biases of Gradient Descent in Offline System Identification and Optimal Control*

Abstract: In control and optimization of critical systems (e.g. systems in healthcare, manufacturing and transportation), trial and error is typically prohibitively costly, prohibitively dangerous, or both. A natural approach for learning to control and optimize without trial and error is using pre-recorded data for offline system identification — i.e. for offline learning of a system model — and then using this model for offline learning of an optimal controller. This approach, namely offline system identification and optimal control, is achieving breakthrough success when implemented with overparameterized models (e.g. neural networks) trained by gradient descent (GD). The success is fueled by implicit biases of GD, which yield not only in-distribution generalization, but also out-of-distribution generalization. Towards elucidating this phenomenon, I will present a series of works that theoretically analyze implicit biases of GD when applied to overparameterized linear models in offline system identification and optimal control. The results I will present offer theoretical explanations for the success of GD in control and optimization of critical systems, and suggest potential avenues for enhancing this success.

12:30 - 14:00 *Lunch* (On your own)

14:00 - 14:30 **Nadav Merlis -** *Stable Matching with Ties: Approximation Ratios and Learning*

Abstract: We study the problem of matching markets with ties, where one side of the market does not necessarily have strict preferences over members on its other side. For example, workers do not always have strict preferences over assignments, students can give the same ranking for different schools and more. In particular, assume w.l.o.g. that workers' preferences are determined by their utility from being matched to each assignment, which might admit ties. Notably, in contrast to classical two-sided markets with strict preferences, there is no longer a single stable matching that simultaneously maximizes the utility for all workers. We aim to guarantee each worker the largest possible share from the utility in her best possible stable matching. We call the ratio between the worker's best possible stable utility and its assigned utility the Optimal Stable Share (OSS)-ratio. We first prove that distributions over stable matchings cannot guarantee an OSS-ratio that is sublinear in the number of workers. Instead, randomizing over possibly non-stable matchings, we show how to achieve a tight logarithmic OSS-ratio. Then, we analyze the case where the real utility is not necessarily known and can only be approximated. In particular, we provide an algorithm that guarantees a similar fraction of the utility compared to the best possible utility. Finally, we move to a bandit setting, where we select a matching at each round and only observe the utilities for matches we perform. We show how to utilize our results for approximate utilities to gracefully interpolate between problems without ties and problems with statistical ties (small suboptimality gaps).

14:30 - 15:00 **Etienne Boursier -** *Mitigating externalities in online learning*

ABSTRACT: In economic theory, the concept of externality refers to any indirect effect resulting from an interaction between players that affects the social welfare. Most of the models within which externality has been studied assume that agents have perfect knowledge of their environment and preferences. This is a major hindrance to the practical implementation of many proposed solutions. To address this issue, we consider a two-player bandit setting where the actions of one of the players affect the other player. We show that the optimal approach for maximizing the social welfare in the presence of externality is to establish property rights, i.e., enable transfers and bargaining between the players. We first show a simple learning algorithm for a single agent, when to other one possesses perfect knowledge of the underlying game. We then design a policy for both players that allows them to learn a bargaining strategy which maximizes the total welfare, recovering the Coase theorem under uncertainty.

15:00 - 15:30 **Tomer Koren -** *Bandits for free in Multiclass Classification*

ABSTRACT: We revisit the classical problem of multiclass classification with bandit feedback (Kakade, Shalev-Shwartz and Tewari, 2008), where each input classifies to one of K possible labels and feedback is restricted to whether the predicted label is correct or not. Our primary inquiry is with regard to the dependency on the number of classes K, which is often very large in multiclass problems. I will survey two recent results in this context: 1. A characterization of the minimax regret of bandit multiclass, establishing that it is of the form $\min |H| + \sqrt{T}, \sqrt{KT \log |H|}$ for finite hypothesis class H; In particular, we present a new bandit classification algorithm that guarantees this rate, improving over classical algorithms (such as EXP4) for moderately-sized hypothesis classes, and give a matching lower bound establishing tightness (up to log-factors) in all parameter regimes. 2. A novel learning algorithm for the agnostic PAC version of the problem, with sample complexity of $O((poly(K) + 1/\epsilon^2) \log(|H|/\delta))$; our algorithm utilizes a stochastic optimization technique to minimize a log-barrier potential based on Frank-Wolfe updates for computing a low-variance exploration distribution over the hypotheses, and is made computationally efficient provided access to an ERM oracle over H. We also provide an extension general classes and establish similar sample complexity bounds in which $\log |H|$ is replaced by the Natarajan dimension. Surprisingly, these results match the asymptotic optimal rates with full-information, and reveal a stark contrast between the PAC and regret-minimization versions of the problem.

Based on joint work with Liad Erez, Alon Cohen, Yishay Mansour and Shay Moran (COLT'24; NeurIPS'24).

15:30 - 16:00 *Coffee Break*

16:00 - 16:30 **Alessandro Lazaric -** *An unsupervised journey in reinforcement learning*

ABSTRACT: Unsupervised reinforcement learning (URL) focuses on how to learn from environment interactions when no explicit reward signal is provided. This setting matches the popular approach of learning from large unsupervised datasets to pre-train foundation models that can be used to solve a wide range of tasks in domains such as computer vision and natural language. In this talk, I'll review a few works on URL ranging from more theoretical insights on autonomous exploration to more algorithmic contributions in learning foundation models using demonstration regularized URL.

16:30 - 17:00 **Claire Vernade -** *Foundations for Continual Reinforcement Learning*

ABSTRACT: Despite its deep connections to dynamical systems, Reinforcement Learning (RL) is typically studied under the assumption of stationarity: the goal is to learn a static policy that maps states to optimal actions. But how well does this assumption hold in real-world applications? Recent advances have led to various attempts to apply RL algorithms

to real-world systems, such as data center cooling or plasma configuration for nuclear fusion reactors. However, these complex systems are often only partially observable and present themselves as non-stationary targets to the controller – a challenge that current RL methods are ill-equipped to handle. In this talk, I will review these challenges and introduce alternative problem definitions designed to make RL more effective in dynamic environments. This work is part of my Emmy Noether project, "Foundations for Lifelong RL." More details are available at www.cvernade.com.

17:00 - 17:30 **Katrina Ligett -** *Doing science: replication, overfitting, and the scientific process*

Abstract: Recent work on the problem of overfitting due to adaptivity in data analysis has made progress on (differential privacy-adjacent) algorithmic tools that can mitigate the risks of adaptivity, formal lower bounds that illustrate bounds on what is achievable in the worst case, and models formalizing adaptivity. However (loosely), the algorithms we have are not practical, the lower bounds we have are not realistic, and the models we have model only a small part of the scientific process. This talk explores directions for future work in all three of these dimensions.