Open Problems: Recent Trends in Algorithms 2024

July 1-3, 2024

The following open problems were discussed in Recent Trends in Algorithms in July 1-3, 2024.

1. Suppose f is a univariate polynomial of degree d with complex coefficients that has exactly s non-zero monomials (i.e. has sparsity s). What can we say about the number of non-zero monomials in f^2 ?

   Clearly, this number is at most $O(s^2)$, and this bound is essentially tight. But what is the best lower bound on the sparsity of f^2, or in general any power of f.

   A classical result of A. Schinzel () proved a bound of $\Omega(\log \log s)$ for this, but as far as I know, the current state of art is still far from well understood.

   Throughout this problem, we should think of s as being much much smaller than d.

2. Suppose $t, \ell \in \mathbb{N}$ such that $t - \ell > 2$ and $S \subseteq \{0, 1, \}^n$ such that $|S| > 1$. What is the minimum degree of a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ such that $\forall u \in S$, P has a zero of multiplicity exactly ℓ at u and $\forall v \in \{0, 1\}^n \setminus S$, P has a zero of multiplicity at least ℓ at v?

   For $|S| = 1$, see [SW20].

3. A set $\mathcal{S} \subset \mathbb{N}$ is called *dissociated* if any equality of the form

   $$\sum_{s \in \mathcal{S}} \varepsilon_s \cdot s = 0$$

   where $\varepsilon_s \in \{-1, 0, +1\}$ implies that all the ε_s's are 0. In other words, all subset sums are distinct.

   It was proven by Richard Guy that for a dissociated set $\mathcal{S}$, we have

   $$|\mathcal{S}(n)| \leqslant \log_2 n + \frac{1}{2} \log_2 \log_2 n + 1.3$$

   where $\mathcal{S}(n) = \mathcal{S} \cap [1, n]$ [Guy82]. It was conjectured by Erdos that the right hand side can be improved to $\log_2 n + c$ for some constant c. Check here for more.

Now, consider the greedy algorithm for Dissociated sets. It starts with two given integers $a > 0$ and $b > a$, and another integer n. It sets $\gamma_1 = a$, $\gamma_2 = b$ and in the r-th step it chooses γ_r to be the smallest integer greater than γ_{r-1} such that the sequence $\{\gamma_k\}_{k=1}^{r}$ is dissociated. The algorithm halts at the n-th step.

Conjecture: For given a, b, let $\Gamma_{a,b} = \{\gamma_1 = a, \gamma_2 = b, \gamma_3, \dots\}$ be the dissociated sequence that the greedy algorithm produces. Then, there is an $n_0 = n_0(a, b)$ such that

$$\gamma_n = 2 \cdot \gamma_{n-1}$$

for all $n \geqslant n_0$.

Moreover, there is an A such that if $a, b > A$, then there is an $n_0 = n_0(a, b)$ such that $\gamma_n < 2 \cdot \gamma_{n-1}$ for $n < n_0$; $\gamma_{n_0} > 2 \cdot \gamma_{n_0-1}$ and $\gamma_n = 2 \cdot \gamma_{n-1}$ for all $n > n_0$.

4. For a graph family $\mathcal{H}$, the $\mathcal{H}$ CONTRACTION problem we are given a graph G and an integer k, and the goal is to check if we can contract at most k edges in G to obtain a graph from $\mathcal{H}$. For hereditary families $\mathcal{H}$, if (G, k) is a yes-instance of $\mathcal{H}$ CONTRACTION then there exists $X \subseteq V(G)$ of size at most $2k$, such that $G - X$ is in $\mathcal{H}$. A generic step one may consider when designing FPT algorithms for the parameter k for a hereditary bipartite family $\mathcal{H}$ is to iterate over the above stated deletion set X gets partitioned into two sets, making the runtime already at least to 4^k. Consider the very simple case when $\mathcal{H}$ is the family of all complete bipartite graphs; the corresponding $\mathcal{H}$ CONTRACTION is called BICLIQUE CONTRACTION. Can you design an algorithm for BICLIQUE CONTRACTION running in time $2^k \cdot |V(G)|^{O(1)}$?

5. Consider the following setup: we are given an $x \in \{-1, 1\}^N$ via an oracle which, for any $i \in [N]$, returns the value of the bit x_i. We also have a real multilinear polynomial p of degree d and N variables with the guarantee that the absolute value of p is bounded (by some constant) for inputs in $\{-1, 1\}^N$. The task is to obtain a good estimate of this polynomial making minimal queries to x. The queries can be adaptive or non-adaptive.

   Let us illustrate this for $d = 1$. Below, we will build an unbiased estimator for p which, on expectation makes *one* (!) oracle query and has a small constant variance.

   Let $p(x_1, \dots, x_N) = \sum_i a_i x_i$. Without loss of generality assume that all the real a_is are non-zero. Also, there exists a constant $c > 0$ such that for all $\widehat{x} \in \{-1, 1\}^N |p(\widehat{x})| \leqslant c$.

   For each $i \in [N]$, define a Bernoulli random variable $\mathbf{X}_i = 1$ with probability $|a_i| / \sum_i |a_i|$ and takes a value 0 otherwise.

   Consider an estimator

$$\mathbf{P}(x) = \sum_{i=1}^{N} |a_i| \left(\sum_{i=1}^{N} sign(a_i) x_i \mathbf{X}_i \right)$$

   In simple terms, this estimator decides to query a variable x_i with probability proportional to its coefficient.

It can be argued (using linearity of expectation) that for any $x \in \{-1, 1\}^N$, $\mathbb{E}[\mathbf{P}(x)] = p(x)$. Moreover, using an application of Khinchine inequality, it can be shown that for all $x \in \{-1, 1\}^N$, its variance $Var(\mathbf{P}) = O(1)$.

Moreover on expectation, the number of variables queried is $\mathbb{E}[\sum_i X_i] = \sum_i \mathbb{E}[X_i] = 1$.

The question is: can we design similar unbiased estimators for p with low number of queries and low variance for values of $d \geqslant 2$ and understand the number of queries versus variance trade-offs in the adaptive and non-adaptive query setting. This problem has applications in distribution distinguishing problem and questions on simulating quantum algorithms classically.

A good starting point on connections to quantum algorithms is a seminal result due to Aaronson and Ambainis [AA18].

For an arbitrary d, if the polynomial corresponds to success probability of a $d/2$ query quantum algorithm (with the string x thought as the oracle), then unbiased estimators are known due to Bravyi *et al.* [BGGS21].

6. What is the **deterministic** query complexity of reconstructing rooted trees with maximum degree $\leqslant d$ using ancestor queries?

   There is a hidden rooted tree with maximum degree $\leqslant d$. You get oracle access to the underlying ancestor-descendant relation in the tree, that is, you can query for any pair (a, b) if a is an ancestor of b. The problem is to reconstruct the tree, minimizing the number of queries. $d = 2$ is a special case of this problem, which corresponds to sorting in the comparison model. [JS13] presents a deterministic algorithm with query complexity $O(dn^{1.5} \log n)$. The best known lower bound is $\Omega(dn \log_d n)$, which also holds for randomized algorithms [BG23], and a randomized algorithm with a matching query and time complexity was presented in [RY23]. In case of deterministic algorithms however, even for rooted binary trees, there is a huge open gap of $\Omega(n \log n)$ vs. $O(n^{1.5} \log n)$.

References

[AA18] Scott Aaronson and Andris Ambainis. Forrelation: A problem that optimally separates quantum from classical computing. *SIAM Journal on Computing*, 47(3):982–1038, 2018.

[BG23] Paul Bastide and Carla Groenland. Optimal distance query reconstruction for graphs without long induced cycles, 2023.

[BGGS21] Sergey Bravyi, David Gosset, Daniel Grier, and Luke Schaeffer. Classical algorithms for forrelation, 2021.

[Guy82] Richard K. Guy. Sets of integers whose subsets have distinct sums. *Annals of Discrete Mathematics*, 12:141–154, 1982.

[JS13] M. Jagadish and Anindya Sen. Learning a bounded-degree tree using separator queries. In Sanjay Jain, Rémi Munos, Frank Stephan, and Thomas

Zeugmann, editors, *Algorithmic Learning Theory*, pages 188–202, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[RY23] Jishnu Roychoudhury and Jatin Yadav. An Optimal Algorithm for Sorting in Trees. In Patricia Bouyer and Srikanth Srinivasan, editors, *43rd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2023)*, volume 284 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:14, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[SW20] Lisa Sauermann and Yuval Wigderson. Polynomials that vanish to high order on most of the hypercube. *arXiv preprint*, 2020.