Tight Dynamic Problem Lower Bounds from Generalized BMM and OMv[*]

Ce Jin[†]
cejin@mit.edu
MIT
Cambridge, MA, USA

Yinzhan Xu[‡]
xyzhan@mit.edu
MIT
Cambridge, MA, USA

Abstract

Popular fine-grained hypotheses have been successful in proving conditional lower bounds for many dynamic problems. Two of the most widely applicable hypotheses in this context are the *combinatorial Boolean Matrix Multiplication (BMM) hypothesis* and the closely-related *Online Matrix Vector Multiplication (OMv) hypothesis*. The main theme of this paper is using k-dimensional generalizations of these two hypotheses to prove new *tight* conditional lower bounds for dynamic problems.

The *combinatorial k-Clique hypothesis*, which is a standard hypothesis in the literature, naturally generalizes the combinatorial BMM hypothesis. In this paper, we prove tight lower bounds for several dynamic problems under the combinatorial k-Clique hypothesis. For instance, we show that.The *Dynamic Range Mode* problem has no combinatorial algorithms with poly(n) pre-processing time, $O(n^{2/3-\epsilon})$ update time and $O(n^{2/3-\epsilon})$ query time for any $\epsilon > 0$, matching the known upper bounds for this problem. Previous lower bounds only ruled out algorithms with $O(n^{1/2-\epsilon})$ update and query time under the OMv hypothesis. The *Dynamic Subgraph Connectivity* problem on undirected graphs with m edges has no combinatorial algorithms with poly(m) pre-processing time, $O(m^{2/3-\epsilon})$ update time and $O(m^{1-\epsilon})$ query time for $\epsilon > 0$, matching the upper bound given by Chan, Pătraşcu, and Roditty [SICOMP'11], and improving the previous update time lower bound (based on OMv) with [a] exponent 1/2Other examples include tight combinatorial lower bounds for *Dynamic 2D Orthogonal Range Color Counting*, *Dynamic 2-Pattern Document Retrieval*, and *Dynamic Range Mode* in higher dimensions.

Furthermore, we propose the OuMv_k hypothesis as a natural generalization of the OMv hypothesis. Under this hypothesis, we prove tight lower bounds for various dynamic problems. For instance, we show that The *Dynamic Skyline Points Counting* problem in $(2k-1)$-dimensional space has no algorithm with poly(n) pre-processing time and $O(n^{1-1/k-\epsilon})$ update and query time for $\epsilon > 0$, even Other examples include tight conditional lower bounds for (semi-online) Dynamic Klee's measure for unit cubes, and high-dimensional generalizations of Erickson's problem and Langerman's problem.

CCS CONCEPTS

- Theory of computation → Data structures design and anal-ysis.

KEYWORDS

fine-grained complexity, dynamic data structures

ACM Reference Format:Ce Jin and Yinzhan Xu. 2022. Tight Dynamic Problem Lower Bounds fromGeneralized BMM and OMv. In Proceedings of the 54th Annual ACM SIGACTSymposium on Theory of Computing (STOC '22), June 20ś24, 2022, Rome, Italy.ACM, New York, NY, USA, 14 pages. `https://doi.org/10.1145/3519935.3520036`

1. INTRODUCTION

Dynamic data structure problems involve maintaining datasets (like graphs or sequences) that undergo small updates while supporting efficient queries.

[a]The full version of this paper is available at https://arxiv.org/abs/2202.11250.

Supported by NSF Grant CCF-2129139.
Supported by NSF Grant CCF-2129139.

This work is licensed under a Creative Commons Attribution 4.0 International License.
STOC '22, June 20–24, 2022, Rome, Italy
©2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9264-8/22/06.
https://doi.org/10.1145/3519935.3520036

These problems are motivated by real-world applications, such as social network graphs and real-time systems, where data frequently changes. Despite progress, many dynamic problems, such as directed graph reachability, still have only slow, polynomial-time solutions. Researchers aim to establish lower bounds for such problems through fine-grained complexity, using hypotheses like combinatorial Boolean Matrix Multiplication (BMM) and Online Matrix-Vector Multiplication (OMv) to prove conditional lower bounds for dynamic problem hardness.

Hypothesis 1.1 (Combinatorial k-Clique Hypothesis). There is no $O(n^{k-\epsilon})$ time combinatorial algorithm for k-Clique Detection on n-vertex graphs, for any $\epsilon > 0$.

Definition 1.2 2 (OuMv_k Problem). During preprocessing we are given a subset $M \subseteq [n]^{k.3}$ Then we receive online queries each specifying k sets $U^{(1)}, U^{(2)}, ..., U^{(k)} \subseteq [n]$, and we need to answer whether $U^{(1)}, U^{(2)}, ..., U^{(k)}$ has a non-empty intersection with M.

Hypothesis 1.3(OuMv_k Hypothesis). There is no algorithm for the OuMv_k problem with n queries in $O(n^{1+k-\epsilon})$ total time (preprocessing time plus total query time) for any $\epsilon > 0$.

Hypothesis 1.4 There is no algorithm for the OuMv_k problem with poly(n) pre-processing time and $O(n^{\gamma+\kappa-\epsilon})$ total query time for n^γ queries, for any $\gamma, \epsilon > 0$.

1.1. *Our Contribution*

Theorem 1.5. Assuming the combinatorial 4-Clique hypothesis, there is no combinatorial data structure that solves Dynamic Range Mode in poly(n) pre-processing time, $O(n^{2/3-\epsilon})$ amortized query time and $O(n^{2/3-\epsilon})$ amortized update time for $\epsilon > 0$. The same lower bound also holds for the Dynamic Range Minority problem.

Previous conditional lower bounds have ruled out algorithms for SubConn with any of the following running times (for any $\epsilon > 0$):

1. (under 3SUM [1]) $O(m^{4/3-\epsilon})$ pre-processing, $O(m^{a-\epsilon})$ update and $O(m^{2/3-a-\epsilon})$ query time, for any $a \in [fraq16, fraq13]$.
2. (under OMv [5]) polynomial pre-processing, $O(m^{a-\epsilon})$ update and $O(m^{1-a-\epsilon})$ query time, for any $a \in (0, 1)$.
3. (under OMv [5]) polynomial pre-processing, $O(m^{1/2-\epsilon})$ update and $O(m^{1-\epsilon})$ query time.

Theorem 1.6 Assuming the combinatorial 4-Clique hypothesis, there is no combinatorial algorithm that solves st-SubConn in (m) pre-processing time, $O(m^{2/3-\epsilon})$ amortized update time, and $O(m^{1-\epsilon})$ amortized query time for $\epsilon > 0$. We leave it as an open problem to either improve the 2/3 exponent in the update time using fast matrix multiplication or determine it's impossible.

Problem 1 [Dynamic 2-Pattern Document Retrieval]. Given a list of strings $S_1, ..., S_D$ of total length $\sum_{i=1}^{D} |S_i| = n$, where each string is on or off, maintain a data structure that supports the following operations:

- Turn on or turn off a string;
- Given a pair of strings (T_1, T_2), count the number of i such that S_i is on and contains both T_1 and T_2.

We show (in the full paper) that this problem can be solved by a combinatorial data structure with $O(n^{2/3})$ time per update and $O(|T_1| + |T_2| + n^{2/3})$ time per query. Under the combinatorial 4-Clique hypothesis, this data structure is in fact optimal among combinatorial ones.

Theorem 1.7 Assuming the combinatorial 4-Clique hypothesis, there is no combinatorial data structure that solves the Dynamic 2-Pattern Document Retrieval problem in (n) pre-processing time, $O(n^{2/3-\epsilon})$ amortized query time and $O(n^{2/3-\epsilon})$ amortized update time for $\epsilon > 0$, even when all patterns have lengths $O(1)$ and the algorithm is only required to determine if the counts are zeros.

Theorem 1.8 Assuming the combinatorial 4-Clique hypothesis, there is no combinatorial data structure that solves Dynamic 2D Orthogonal Range Color Counting in (n) pre-processing time, $O(n^{2/3-\epsilon})$ amortized query time and $O(n^{2/3-\epsilon})$ amortized update time for $\epsilon > 0$.

Theorem 1.9 Assuming the combinatorial 4-Clique hypothesis, there is no combinatorial data structure that solves *st*-Reach, Dynamic Strong Connectivity, or Dynamic Bipartite Perfect Matching, in (n) pre-processing time, $O(n^{2-\epsilon})$ amortized query time and $O(n^{2-\epsilon})$ amortized update time for $\epsilon > 0$.

Theorem 1.10 Let $k \geq 2$ be a positive integer. Assuming the OuMv_k hypothesis, there is no data structure for Dynamic Skyline Points Counting in $\mathbb{R}^{2k-1}$ with (n) pre-processing time, $O(n^{1-1/k-\epsilon})$ amortized update and query time for $\epsilon > 0$, even in the semi-online model.

Theorem 1.11 Let $k \geq 2$ be a positive integer. Assuming the OuMv_k hypothesis, there is no data structure for Dynamic Klee's measure for unit hypercubes in $\mathbb{R}^{2k-1}$ with (n) pre-processing time, $O(n^{1-1/k-\epsilon})$ amortized update and query time for $\epsilon > 0$, even in the semi-online model.

Theorem 1.12 Let $k \geq 2$ be a positive integer. Assuming the OuMv$_k$ hypothesis, there is no data structure for Chan's Halfspace problem in $\mathbb{R}^k$ with (n) pre-processing time, $O(n^{1-1/k-\epsilon})$ amortized update and query time for $\epsilon > 0$, even if we only need to output $\min c_H(Q)$ for each query.

Table 1: Our lower bounds for dynamic problems. The lower bounds based on the k-Clique hypothesis work for combinatorial algorithms and the lower bounds based on the OuMv$_k$ hypothesis work for arbitrary algorithms. The lower bounds state that there are no algorithms achieving the stated pre-processing time, update time and query time simultaneously for $\epsilon > 0$, under the corresponding hypotheses, even if the algorithms have amortized update and query time, and the updates are semi-online. All our lower bounds have matching upper bounds unless otherwise stated. The upper bounds column references algorithms that run in the stated pre-processing time, update time and query time simultaneously for $\epsilon = 0$, up to poly-logarithmic factors. All algorithms work for fully dynamic inputs with worst-case time guarantees, unless otherwise stated.

Problems	Lower Bounds			Hypotheses	References	Upper Bounds
	Pre-processing	Update	Query			
Dynamic Range Mode	poly(n)	$n^{2/3-\epsilon}$	$n^{2/3-\epsilon}$	4-Clique	Thm. 1.5	[[2],[8]]
Dynamic Range Minority	poly(n)	$n^{2/3-\epsilon}$	$n^{2/3-\epsilon}$	4-Clique	Thm. 1.5	[8]
Dynamic d-dimensional Orthogonal Range Mode	poly(n)	$n^{1-1/(2d+1)-\epsilon}$	$n^{1-1/(2d+1)-\epsilon}$	(2d+2)-Clique	Thm. 3.3	Prop. 3.2
st Subgraph Connectivity	poly(n)	$m^{2/3-\epsilon}$	$m^{1-\epsilon}$	4-Clique	Thm. 1.6	[4] amortized
Dynamic 2-Pattern Document Retrieval	poly(n)	$n^{2/3-\epsilon}$	$n^{2/3-\epsilon}$	4-Clique	Thm. 1.7	Full Paper
Dynamic 2D Orthogonal Range Color Counting	poly(n)	$n^{2/3-\epsilon}$	$n^{2/3-\epsilon}$	4-Clique	Thm. 1.8	Full Paper
Dynamic st-Reachability	poly(n)	$n^{2-\epsilon}$		4-Clique	Thm. 1.9	trivial
Dynamic Strong Connectivity						
Dynamic Bipartite Perfect Matching						
Dynamic Skyline Points Counting in $\mathbb{R}^{2k-1}$	poly(n)	$n^{1-1/k-\epsilon}$	$n^{1-1/k-\epsilon}$	OuMv$_k$	Thm. 1.10	Prop. 4.4 semi-online
Dynamic Klee's Measure for unit hypercubes in $\mathbb{R}^{2k-1}$	poly(n)	$n^{1-1/k-\epsilon}$	$n^{1-1/k-\epsilon}$	OuMv$_k$	Thm. 1.11	[3] semi-online only for $k = 2$
Clan's Halfspace problem in $\mathbb{R}^k$	poly(n)	$n^{1-1/k-\epsilon}$	$n^{1-1/k-\epsilon}$	OuMv$_k$	Thm. 1.12	[3] amortized
Dynamic s-k-Uniform (k+1)-Hyperclique	poly(n)	$n^{1-\epsilon}$	$n^{k-\epsilon}$	OuMv$_k$	Full Paper	trivial
k-Dimensional Erickson's problem	poly(n)	$n^{1-\epsilon}$	$n^{k-\epsilon}$	OuMv$_k$	Full Paper	trivial
(k-1)-Dimensional Langerman's problem	poly(n)	$n^{(k-1)^2/k-\epsilon}$	$n^{(k-1)^2/k-\epsilon}$	OuMv$_k$	Full Paper	Full Paper

1.2. Further Related Works

Pătraşcu [7] was arguably the first to systematically study finegrained conditional lower bounds for dynamic problems. In this groundbreaking work, Pătraşcu first reduced the 3SUM problem to some triangle reporting problem, which is then further reduced to many dynamic problems such as Dynamic Reachability, Dynamic Shortest Paths and Subgraph Connectivity. This series of reductions show polynomial lower bounds for dynamic problems under the 3SUM hypothesis. This work was later generalized by, for instance, Abboud and Vassilevska Williams [1], and Kopelowitz, Pettie, and Porat [6] to show polynomial lower bounds for more problems under the 3SUM hypothesis. Both [1] and [6] use some variants of the aforementioned triangle reporting problem as intermediate steps in their reductions.

2. PRELIMINARIES

In a graph G = (V , E), we use $N(v)$ to denote the set of neighbors of $\sqsubseteq \in V$. For any subset $U \subseteq V$, we use $\mathcal{N}_{\sqcap}(v)$ to denote N $\mathcal{N}_{\sqcap}(v) \cap U$. By known techniques (e.g. [92]), the combinatorial k-Clique hypothesis is equivalent to the following unbalanced version. Hypothesis 2.1 (Combinatorial k-Clique Hypothesis, unbalanced version). Let $d_1, d_2, ..., d_k > 0$ be constant real numbers. There is no $O(n^{d_1+d_2+...+d_k-\epsilon})$-time combinatorial algorithm for k-Clique on k-partite graphs $(V_1 \cup V_2 \cup .. \cup V_k, E) where |V_i| = n^{d_i} for i \in [k], for any \epsilon > 0$

3. LOWER BOUNDS UNDER THE k-CLIQUE HYPOTHESIS

In this section, we show tight combinatorial lower bounds for Dynamic Range Mode and st Subgraph Connectivity under the combinatorial 4-Clique hypothesis. We defer the proofs for Dynamic 2-Pattern Document Retrieval and Dynamic 2D Orthogonal Range Color Counting to the full paper.

3.1. Range Mode

We first recall the definition of Dynamic Range Mode.

Problem 2 [Dynamic Range Mode] Maintain a data structure for an integer array a of size at most n, and support the following operations:

- Insert or delete an integer;
- For each query specified by l, r, report the most frequent integer appearing in $a_l, a_{l+1}, \ldots, a_r$, breaking ties arbitrarily.

Theorem 1.5 Assuming the combinatorial 4-Clique hypothesis, there is no combinatorial data structure that solves Dynamic Range Mode in poly(n) pre-processing time, $O(n^{2/3-\epsilon})$ amortized query time and $O(n^{2/3-\epsilon})$ amortized update time for $\epsilon > 0$. The same lower bound also holds for the Dynamic Range Minority problem.

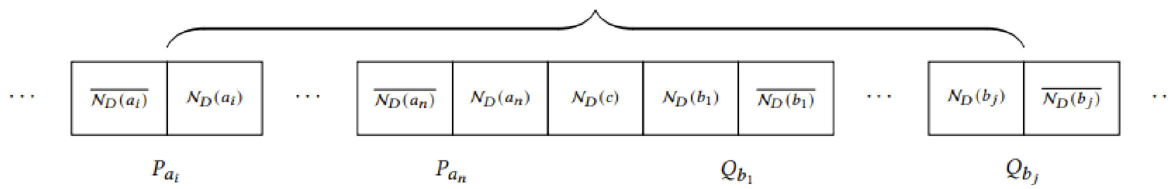

Figure 1: This figure depicts the range mode query corresponding to vertices a_i, b_j and c. Here, $\mathcal{N}_D(v)$ denotes the set of neighbors of vertex v in D, and $\overline{\mathcal{N}_D(v)}$ denotes the set of non-neighbors of vertex v in D.

Theorem 3.1 Assuming the $(2d+1)$-Clique hypothesis, there is no combinatorial data structure that solves Batch d-Dimensional Orthogonal Range Mode in $O(n^{2-1/2d-\epsilon})$ time for $\epsilon > 0$.

PROOF. We reduce from an unbalanced instance of $(2d+1)$-Clique Detection, where the first $2d$ parts $V_1, ..., V_{2d}$ all have sizes $n^{1/2d}$, while the last part V_{2d+1} has size $n^{1-1/2d}$. By Fact, combinatorial algorithms for such unbalanced instances of $(2d+1)$-Clique Detection require $n^{2-1/2d-o(1)}$ time under the combinatorial $(2d+1)$-Clique hypothesis.

For each $i \in [2d]$, we will first create an array A_i of size $O(n)$ as follows. The elements in the array will be identified by vertices in V_{2d+1}. For each $v_i \in V_i$, we create a permutation of V_{2d+1}, such that the neighbors of v_i in V_{2d+1} appear before the non-neighbors of v_i in V_{2d+1}. The array A_i is then the concatenation of all the permutations.

We can split each of the d axes in R^d at the origin to get a total of $2d$ half-axes. We will put each A_i on one of the half-axes as follows. For each odd $i \in [2d]$ and each $j \in [|A_i|]$, we add a point whose $\lceil i/2 \rceil$-th coordinate is j and whose other coordinates are all zeros. We assign a label $A_i[j]$ to this point. For each even $i \in [2d]$ and each $j \in [|A_i|]$, we add a point whose $\lceil i/2 \rceil$-th coordinate is $-j$ and whose other coordinates are all zeros. We similarly assign a label $A_i[j]$ to this point.

Fix a tuple $(v_1, ..., v_{2d}) \in V_1 \times ... \times V_{2d}$. For every $i \in [2d]$, we use b_i to denote the index in A_i of the last neighbor of v_i in the permutation corresponding to $\mathcal{N}_{V_{2d+1}}(v_i)$. Then we ask a range mode query on the orthogonal range defined as the following:

$$\begin{cases} x_{\lceil (i/2) \rceil} \le b_i : i \in [2d] & \text{is odd,} \\ x_{\lceil (i/2) \rceil} \ge -b_i : i \in [2d] & \text{is even.} \end{cases}$$

It is not hard to see that the multi-set of labels in this orthogonal range is exactly

$$\{A_i[j] : 1 \le i \le 2d, 1 \le j \le b_i\}.$$

By construction of the arrays A_i, this multiset is the union of several full permutations of V_{2d+1}, and the neighborhoods of $v_1, ..., v_{2d}$ in V_{2d+1}. Thus, if $v_1, ..., v_{2d}$ have a common neighbor in V_{2d+1}, the mode of the orthogonal range will also be a common neighbor.

Therefore, by asking $O(n)$ range mode queries on this instance, we are able to determine whether each tuple $(v_1, ..., v_{2d}) \in V_1 \times \cdots \times V_{2d}$ has a common neighbor in V_{2d+1}, so that we can solve the $(2d+1)$-Clique

Detection instance in $O(n)$ additional time. This concludes the lower bound proof for Batch d-Dimensional Orthogonal Range Mode.

Similar to Dynamic Range Mode, the Batch d-Dimensional Orthogonal Range Mode has a natural dynamic variant.

Proposition 3.2 There is a combinatorial data structure that solves Dynamic d-Dimensional Orthogonal Range Mode in $O(n^{2-2/(2d+1)})$ pre-processing time, $O(n^{1-1/(2d+1)})$ query time and $O(n^{1-1/(2d+1)})$ update time.

Theorem 3.3 Assuming the combinatorial $(2d+2)$-Clique hypothesis, there is no combinatorial data structure that solves Dynamic d-Dimensional Orthogonal Range Mode in (n) pre-processing time, $O(n^{1-1/(2d+1)-\epsilon})$ amortized query time and $O(n^{1-1/(2d+1)-\epsilon})$ amortized update time for $\epsilon > 0$.

3.2. st Subgraph Connectivity

Algorithm 1: the reduction from 4-Clique Detection to st-SubConn

Initialize the st-SubConn data structure on G,letting S contain all vertices.
for $d \in D$ **do**
  for $a \in A$ **do**
    Let $a^{U_D} \in S$ if and only if $(a, d) \in E$.
  for $b \in B$ such that $(b, d) \in E$ **do**
    Let $b^{V_B} \in S$.
    for $b' \in B/\{b\}$ **do**
      Let $(b')^{V_B} \notin S$
    for $c \in C$ **do**
      Let $c^{V_c} \in S$ if and only if $(c, d), (c, b) \in E$
    if s,t are connected in the induced subgraph of S **then**
      return True
return False

4. GEOMETRIC PROBLEMS AND OuMv$_k$ HYPOTHESIS

4.1. Skyline Points Counting

Proposition 4.4 For any $k \geq 2$, there exists a data structure for Dynamic Skyline Points Counting in R^{2k-1} in the semi-online model with (n) preprocessing time and $(n^{1-1/k})$ update and query time.

ACKNOWLEDGMENTS

We would like to thank Virginia Vassilevska Williams for many helpful discussions during the early phase of this project. We also thank her for valuable comments on a draft of this paper.

References

[1] Amir Abboud and Virginia Vassilevska Williams. "Popular Conjectures Imply Strong Lower Bounds for Dynamic Problems". In: *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*. 2014, pp. 434–443. DOI: 10.1109/FOCS.2014.53.

[2] Timothy Chan et al. "Linear-Space Data Structures for Range Mode Query in Arrays". In: vol. 55. May 2012, pp. 290–301. ISBN: 978-3-642-31154-3. DOI: 10.1007/s00224-013-9455-2.

[3] Timothy M. Chan. "Semi-Online Maintenance of Geometric Optima and Measures". In: *SIAM Journal on Computing* 32.3 (2003), pp. 700–716. DOI: 10.1137/S0097539702404389. URL: https://doi.org/10.1137/S0097539702404389.

[4] Timothy M. Chan, Mihai Pa˘traşcu, and Liam Roditty. "Dynamic Connectivity: Connecting to Networks and Geometry". In: *SIAM Journal on Computing* 40.2 (2011), pp. 333–349. DOI: 10.1137/090751670.

[5] Monika Henzinger et al. "Unifying and Strengthening Hardness for Dynamic Problems via the Online Matrix-Vector Multiplication Conjecture". In: *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*. STOC '15. Portland, Oregon, USA: Association for Computing Machinery, 2015, pp. 21–30. ISBN: 9781450335362. DOI: 10.1145/2746539.2746609. URL: https://doi.org/10.1145/2746539.2746609.

[6] Tsvi Kopelowitz, Seth Pettie, and Ely Porat. "Higher Lower Bounds from the 3SUM Conjecture". In: *Proceedings of the 2016 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 1272–1287. DOI: 10.1137/1.9781611974331.ch89. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611974331.ch89. URL: https://epubs.siam.org/doi/abs/10.1137/1.9781611974331.ch89.

[7] Mihai Patrascu. "Towards polynomial lower bounds for dynamic problems". In: *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*. STOC '10. Cambridge, Massachusetts, USA: Association for Computing Machinery, 2010, pp. 603–610. ISBN: 9781450300506. DOI: 10.1145/1806689.1806772. URL: https://doi.org/10.1145/1806689.1806772.

[8] Hicham El-Zein et al. "Improved Time and Space Bounds for Dynamic Range Mode". In: *26th Annual European Symposium on Algorithms (ESA 2018)*. Ed. by Yossi Azar, Hannah Bast, and Grzegorz Herman. Vol. 112. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018, 25:1–25:13. ISBN: 978-3-95977-081-1. DOI: 10.4230/LIPIcs.ESA.2018.25. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.ESA.2018.25.