

Space-Efficient Fault-Tolerant Diameter Oracles

Davide Bilò  

Department of Humanities and Social Sciences, University of Sassari, Italy

Sarel Cohen 

Hasso Plattner Institute, University of Potsdam, Germany

Tobias Friedrich  

Hasso Plattner Institute, University of Potsdam, Germany

Martin Schirneck 

Hasso Plattner Institute, University of Potsdam, Germany

Abstract

We design f -edge fault-tolerant diameter oracles (f -FDO, or simply FDO if $f = 1$). For a given directed or undirected and possibly edge-weighted graph G with n vertices and m edges and a positive integer f , we preprocess the graph and construct a data structure that, when queried with a set F of edges, where $|F| \leq f$, returns the diameter of $G - F$. An f -FDO has stretch $\sigma \geq 1$ if the returned value \hat{D} satisfies $\text{diam}(G - F) \leq \hat{D} \leq \sigma \text{diam}(G - F)$.

For the case of a single edge failure ($f = 1$) in an unweighted directed graph, there exists an approximate FDO by Henzinger et al. [ITCS 2017] with stretch $(1 + \varepsilon)$, constant query time, space $O(m)$, and a combinatorial preprocessing time of $\tilde{O}(mn + n^{1.5} \sqrt{Dm/\varepsilon})$, where D is the diameter.

We present an FDO for directed graphs with the same stretch, query time, and space. It has a preprocessing time of $\tilde{O}(mn + n^2/\varepsilon)$, which is better for constant $\varepsilon > 0$. The preprocessing time nearly matches a conditional lower bound for combinatorial algorithms, also by Henzinger et al. With fast matrix multiplication, we achieve a preprocessing time of $\tilde{O}(n^{2.5794} + n^2/\varepsilon)$. We further prove an information-theoretic lower bound showing that any FDO with stretch better than $3/2$ requires $\Omega(m)$ bits of space. Thus, for constant $0 < \varepsilon < 3/2$, our combinatorial $(1 + \varepsilon)$ -approximate FDO is near-optimal in all parameters.

In the case of multiple edge failures ($f > 1$) in undirected graphs with non-negative edge weights, we give an f -FDO with stretch $(f + 2)$, query time $O(f^2 \log^2 n)$, $\tilde{O}(fn)$ space, and preprocessing time $\tilde{O}(fm)$. We complement this with a lower bound excluding any finite stretch in $o(fn)$ space.

Many real-world networks have polylogarithmic diameter. We show that for those graphs and up to $f = o(\log n / \log \log n)$ failures one can swap approximation for query time and space. We present an exact combinatorial f -FDO with preprocessing time $mn^{1+o(1)}$, query time $n^{o(1)}$, and space $n^{2+o(1)}$. When using fast matrix multiplication instead, the preprocessing time can be improved to $n^{\omega+o(1)}$, where $\omega < 2.373$ is the matrix multiplication exponent.

2012 ACM Subject Classification Theory of computation \rightarrow Shortest paths; Theory of computation \rightarrow Data structures design and analysis; Theory of computation \rightarrow Cell probe models and lower bounds; Theory of computation \rightarrow Pseudorandomness and derandomization

Keywords and phrases derandomization, diameter, distance sensitivity oracle, fault-tolerant data structure, space lower bound

Digital Object Identifier 10.4230/LIPIcs.MFCS.2021.18

Related Version *Full Version:* <https://arxiv.org/abs/2107.03485>

Funding *Davide Bilò:* This work was partially supported by the Research Grant FBS2016_BILO, funded by “Fondazione di Sardegna” in 2016.



© Davide Bilò, Sarel Cohen, Tobias Friedrich, and Martin Schirneck;
licensed under Creative Commons License CC-BY 4.0

46th International Symposium on Mathematical Foundations of Computer Science (MFCS 2021).

Editors: Filippo Bonchi and Simon J. Puglisi; Article No. 18; pp. 18:1–18:16

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

The diameter is one of the most fundamental graph parameters. It plays a particular significant role in the analysis of communication networks as the time to transmit a message to all nodes is strongly related with the diameter. Several lines of work have recently attacked the problem of computing the diameter in different settings. For example, Choudhary and Gold [16] constructed diameter spanners, which are subgraphs that approximately preserve the diameter of the original graph, Ancona et al. [6] developed algorithms for computing the diameter in dynamic scenarios and proved matching conditional lower bounds, and Bonnet [10] proved that, for any constant $\varepsilon > 0$, computing a $(7/4 - \varepsilon)$ -approximation of the diameter of a sparse graph n vertices and $m = n^{1+o(1)}$ edges requires $m^{4/3-o(1)}$ time, unless the Strong Exponential Time Hypothesis fails.

In this paper, we approach the diameter from the perspective of fault tolerance. A communication network may be subject to a small number of transient failures, and we want to quickly find out the new diameter without recomputing it from scratch. Therefore, we study the problem of constructing space-efficient data structures that can quickly report the diameter even if up to f edges fail in the graph. We refer to them as *f-edge fault-tolerant diameter oracles* (f -FDO, or simply FDO if $f = 1$). More precisely, given an undirected or directed and possibly edge-weighted graph G and a positive integer f , we want to construct an f -FDO that, when queried on a set F of up to f edges of G , returns a value \widehat{D} that is always at least as large the diameter of $G - F$, denoted by $\text{diam}(G - F)$. We say that an f -FDO has a *stretch* of $\sigma \geq 1$ (or that it is *σ -approximate*) if the value \widehat{D} returned by the oracle additionally satisfies $\text{diam}(G - F) \leq \widehat{D} \leq \sigma \text{diam}(G - F)$.

When designing f -FDOs one must find a good compromise between the following parameters: the stretch, the time needed to query the oracle, the size of the data structure, and the preprocessing time needed to build it. We focus particularly on space-optimal solutions, while keeping the query and preprocessing times low. For the case of a single edge failure in undirected edge-weighted graphs, there are two folklore FDOs known. One reports the exact diameter and has size $O(m)$, while the other takes $O(n)$ space, but guarantees only a stretch of 2. (more details are given in Section 2.1.) In a sense they mark the extreme points of a spectrum. It is natural to ask whether there are more trade-offs possible between the stretch and size of an FDO. More precisely, we pose the following question.

Question 1 – space vs. approximation trade-off. What is the minimum achievable size of an FDO for a given stretch σ ? To answer the question, we prove an information-theoretic lower bound. It shows that for undirected unweighted graphs and every (even non-constant) $1 \leq \sigma < 3/2$, every σ -approximate diameter oracle requires $\Omega(m)$ bits of space. The space lower bound also holds for the harder case of directed graphs. The size of the exact folklore FDO is thus optimal up to the size of a machine word. Moreover, we prove that the stretch 2 of the approximate FDO cannot be improved on weighted graphs while keeping $O(n)$ space.

► **Theorem 1.** *Any FDO with stretch $\sigma = \sigma(m) < 3/2$ must take $\Omega(m)$ bits of space on undirected graphs with m edges. The bound increases to $\sigma < 2$ if the graphs are edge-weighted.*

When we focus our attention on the preprocessing time, the exact FDO can be constructed in $\widetilde{O}(n^3)$ time¹ using the distance sensitivity oracle (DSO) of Bernstein and Karger [8]. Henzinger et al. [30] proved an essentially matching conditional lower bound for combinatorial²

¹ For a positive function $g(n, m, f)$, we use $\widetilde{O}(g)$ to denote $O(g \cdot \text{polylog}(n))$.

² The term “combinatorial algorithm” is not well-defined, and is often interpreted as not using any matrix multiplication. Arguably, combinatorial algorithms can be considered efficient in practice as the constants hidden in the matrix multiplication bounds are rather high.

algorithms. They assumed that any combinatorial algorithm requires $n^{3-o(1)}$ time to multiply two Boolean $n \times n$ matrices, known as the BMM conjecture. The restriction to combinatorial algorithms is crucial as the task is reducible to integer matrices and one can use fast matrix multiplication to solve it in $O(n^\omega)$ time, where $\omega < 2.37286$ is the matrix multiplication exponent [2]. Under the BMM conjecture, Henzinger et al. [30] showed that, for $0 < \varepsilon < 1/3$, any combinatorial preprocessing algorithm requires $n^{3-o(1)}$ time to build an FDO of stretch $(1 + \varepsilon)$, even if we allow $O(n^{2-\delta})$ query time for any constant $\delta > 0$.

They match this bound with an FDO with stretch $(1 + \varepsilon)$ and $O(1)$ query time that can be constructed in time $\tilde{O}(mn + n^{1.5} \sqrt{\text{diam}(G)} \cdot m/\varepsilon)$. Their oracle also reports the radius and vertex eccentricities in the presence of a single edge failure. Even on sparse graphs with $m = \tilde{O}(n)$ edges and constant diameter, the preprocessing time is $\tilde{O}(n^{2.5}/\sqrt{\varepsilon})$. For constant $\varepsilon > 0$, this is by a factor \sqrt{n} larger than the $\tilde{O}(mn)$ time needed to build the DSO of Bernstein and Karger [8]. It is interesting whether one can close the gap.

Question 2 – fast preprocessing time. Does there exist a combinatorial algorithm that constructs in $\tilde{O}(mn)$ time an FDO with stretch $(1 + \varepsilon)$ and constant query time? In addition, can one bypass the combinatorial lower bound by using fast matrix multiplication? We answer these questions affirmatively for the diameter case with the following theorem. The proof of the algebraic part uses the DSO presented very recently by Gu and Ren [28].

► **Theorem 2.** *For every unweighted directed graph and $\varepsilon > 0$, there exists a randomized combinatorial $(1 + \varepsilon)$ -approximate FDO that takes $O(m)$ space and has $\tilde{O}(mn + n^2/\varepsilon)$ preprocessing time and $O(1)$ query time. The returned values are correct w.h.p.³ Using fast matrix multiplication instead, one can construct the FDO in time $\tilde{O}(n^{2.5794} + n^2/\varepsilon)$.*

Note that, for any constant $0 < \varepsilon < 1/3$, our combinatorial $(1 + \varepsilon)$ -approximate combinatorial FDO from Theorem 2 is near-optimal with respect to all parameters. The $\Theta(m)$ space is near-optimal by Theorem 1, the query time is $\tilde{O}(1)$, and the $\tilde{O}(mn)$ preprocessing time comes within sub-polynomial factors of the conditional lower bound by Henzinger et al. [30]. Furthermore, when fast matrix multiplication is permitted, our algebraic preprocessing algorithm is even faster on dense graphs. However, our FDO is randomized.

Question 3 – derandomization. Can the construction of Theorem 2 be derandomized in the same asymptotic running time? We answer this question partially in that we derandomize the approximation part of our algorithm. When combined with the DSO of Bernstein and Karger [8] this gives a deterministic combinatorial FDO. For the derandomization, we adapt the framework of Alon, Chechik, and Cohen [3]. We identify a set of $O(n^{3/2})$ critical paths one needs to hit, and show how to compute them in $O(mn)$ time. It is then enough to let the folklore greedy algorithm compute a hitting set in $\tilde{O}(n^2)$ time.

It remains an open problem whether one can derandomize the algebraic approach, whose randomization stems solely from the DSO by Gu and Ren [28].

► **Theorem 3.** *For every unweighted directed graph and $\varepsilon > 0$, there exists a deterministic combinatorial $(1 + \varepsilon)$ -approximate FDO that takes $O(m)$ space and has $\tilde{O}(mn + n^2/\varepsilon)$ preprocessing time and $O(1)$ query time.*

³ An event occurs *with high probability* (w.h.p.) if it has probability at least $1 - n^{-c}$ for some $c > 0$.

Question 4 – space vs. approximation trade-off for multiple failures. Finally, we consider the case of multiple edge failures and examine similar questions. What is the minimum size for an exact, respectively, approximate, diameter oracle in the presence of up to f edge failures? We again prove an information-theoretic lower bound and show that for arbitrary finite stretch σ , any σ -approximation diameter oracle requires $\Omega(fn)$ bits of space, at least if the oracle can be queried also with sets F that contain non-edges.

► **Theorem 4.** *Suppose $f < n$. Any f -FDO with finite stretch that can be queried also for non-edges must take $\Omega(fn)$ bits of space on graphs with n vertices.*

We develop an efficient f -FDO whose space requirement almost matches the lower bound. Our result adapts and improves a construction by Bilò et al. [9]. Note that we use the \tilde{O} -notation to suppress polylogarithmic factors in n .

► **Theorem 5.** *For every undirected graph with non-negative edge weights, there exists a deterministic combinatorial $(f + 2)$ -approximate f -FDO that takes $\tilde{O}(fn)$ space and has $\tilde{O}(fm)$ preprocessing time and $\tilde{O}(f^2)$ query time.*

Real-world networks are often described as having a small diameter, dubbed as the “small world property” [35]. Many graph models used to analyze social and communication networks have provable polylogarithmic guarantees on the diameter, e.g. Chung-Lu graphs [17], hyperbolic random graphs [24], or the preferential attachment model [31]. We show that on graphs with low diameter one can swap approximation for query time even for multiple failures, while still retaining efficient preprocessing time and a low space requirement. To achieve this, we combine fault-tolerant trees that were introduced by Chechik et al. [13] with the random graphs of Weimann and Yuster [43].

► **Theorem 6.** *Let f be a positive integer and $\delta = \delta(n, m) > 0$ a real number. For every undirected unweighted graph with diameter at most $n^{\delta/f}/(f+1)$, there exists a randomized combinatorial f -FDO that takes $O(n^{2+\delta})$ space, has $O(2^f)$ query time, and with high probability $\tilde{O}(fmn^{1+\delta} + fn^{2+(2-1/f)\delta})$ preprocessing time. Using fast matrix multiplication instead, one can construct the FDO w.h.p. in time $\tilde{O}(fn^{\omega+\delta} + fn^{2+(2-1/f)\delta})$.*

If the diameter is in fact polylogarithmic and the number of failures is bounded by $f = o(\log n / \log \log n)$, we obtain the following corollary.

► **Corollary 7.** *Let $f = o(\log n / \log \log n)$. For every undirected graph with polylogarithmic diameter, there is an f -FDO that takes $n^{2+o(1)}$ space and has $n^{o(1)}$ query time. It can be preprocessed in time $mn^{1+o(1)}$, or algebraically in time $n^{\omega+o(1)}$. If f is constant, the preprocessing times are $\tilde{O}(mn)$, resp. $\tilde{O}(n^\omega)$, with $\tilde{O}(n^2)$ space and $\tilde{O}(1)$ query time.*

1.1 Related Work

We briefly review previous work on distance sensitivity oracles and diameter computation.

Distance sensitivity oracles. Distance oracles for all-pairs distances were introduced in a seminal paper by Thorup and Zwick [42]. Demetrescu et al. [19] extended the notion of distance oracles to the fault-tolerant setting in which either an edge or a vertex of a graph can fail (i.e., distance sensitivity oracles or DSOs). They showed that it is possible to preprocess a directed weighted graph in $\tilde{O}(mn^2)$ time to compute a data-structure of size $O(n^2 \log n)$ capable of answering distance queries in constant time. Bernstein and Karger [8] improved the preprocessing time to $\tilde{O}(mn)$ and Duan and Zhang [23] reduced the space to $O(n^2)$, which is asymptotically optimal.

Duan and Pettie [22] considered the more involved case of two failures and presented an oracle with $O(n^2 \log^3 n)$ size, $O(\log n)$ query time and polynomial construction time. Chechik et al. [13] presented a DSO of size $O(n^{2+o(1)})$ that supports up to $o(\log n / \log \log n)$ edge failures and guarantees a stretch of $(1 + \epsilon)$, for every constant $\epsilon > 0$. The approach has been recently extended to also handle vertex failures by Duan, Gu, and Ren [21].

The construction of DSOs have also been considered in the approximate regime [13]. Algebraic algorithms are known to improve the preprocessing times, if one is willing to employ fast matrix multiplication (for e.g., see [28, 12] and the references therein).

Diameter computation. The fastest known combinatorial algorithms (up to polylogarithmic factor) for both solving the all-pairs shortest paths (APSP) problem and the diameter problem, are the trivial ones with $\tilde{O}(mn)$ running time. There is extensive research on developing faster approximate APSP algorithms [7, 18, 33], as well as faster approximation algorithms for the diameter [14, 40]. For special classes of graphs, for example planar graphs, efficient exact algorithms for computing the diameter are known [25].

2 Preliminaries

We let $G = (V, E)$ denote the (possibly directed) base graph on n vertices and m edges. We tacitly assume that G is (strongly) connected, i.e., $m = \Omega(n)$. For a graph H , we denote by $V(H)$ the set of its vertices, and by $E(H)$ its edges. The (*closed*) *neighborhood* of a vertex $v \in V(H)$ is the set $N[v] = \{u \in V(H) \mid \{v, u\} \in E(H)\} \cup \{v\}$. Let P be a path in H , its *length* $|P|$ is the number of its edges. For any two vertices $x, y \in V(P)$, $P[x..y]$ is the subpath of P from x to y . For $s, t \in V(H)$, the *distance* $d_H(s, t)$ is the minimum length of an s - t -paths in H ; if s and t are disconnected, we set $d_H(s, t) = +\infty$. We drop the subscript when talking about the base graph G . The *eccentricity* of s is $\text{ecc}(s, H) = \max_{t \in V(H)} d_H(s, t)$ and the *diameter* is $\text{diam}(H) = \max_{s \in V(H)} \text{ecc}(s, H)$. Any graph distance can be stored in a single machine word on $O(\log n)$ bits. Unless explicitly stated otherwise, we measure the space complexity in the number of words. For a collection $F \subseteq \binom{V(H)}{2}$ of 2-sets of vertices (edges or non-edges), let $H - F$ be the graph obtained from H by removing all edges in F (graph H is not altered if $F \cap E(H) = \emptyset$). A *replacement path* $P_H(s, t, F)$ is a shortest path from s to t in $H - F$. Its length $d_H(s, t, F) = |P_H(s, t, F)|$ is the *replacement distance*. The *fault-tolerant diameter* of H with respect to F is the diameter of $H - F$.

For a positive integer f , an f -*fault-tolerant diameter oracle* (f -FDO) for the graph G is a data structure that reports, upon query F with $|F| \leq f$, the value $\text{diam}(G - F)$. For any $\sigma = \sigma(n, m, f) \geq 1$, such an oracle is σ -*approximate*, or has *stretch* σ , if it answers a query F with a value \hat{D} such that $\text{diam}(G - F) \leq \hat{D} \leq \sigma \cdot \text{diam}(G - F)$. In case of a single failure, we write FDO for 1-FDO and abbreviate $F = \{e\}$ to e . An f -*distance sensitivity oracle* (f -DSO) reports, upon query (s, t, F) with $|F| \leq f$, the replacement distance $d(s, t, F)$.

2.1 (Mostly) Known FDOs for Single Edge Failures

The first folklore FDO handles single edge failures in unweighted (directed or undirected) graphs. It has also been observed in [30]. The DSO of Bernstein and Karger [8] constructible in $\tilde{O}(mn)$ time and is able to report in constant time the exact distance of any pair of vertices in the presence of a single edge failure. With this one can construct the FDO by explicitly computing all the eccentricities $\text{ecc}(v, G - e)$, for every vertex v and every edge e of G , in $O(n^3)$ time. For a fixed vertex v , the m values $\text{ecc}(v, G - e)$ can be obtained in $O(n^2)$ time as follows. First compute a shortest paths tree T_v of G rooted at v . For each edge e that

is not in T_v , we have that $\text{ecc}(v, G - e) = \text{ecc}(v, G)$. For the tree-edges e in T_v , we use the DSO to compute $\text{ecc}(v, G - e)$ which is the maximum distance from v to any other vertex in $G - e$. Therefore, $\text{ecc}(v, G - e)$ can be computed by performing $n - 1$ queries, as there are $n - 1$ edges in T_v , we need $O(n^2)$ time. The fault-tolerant diameter $\text{diam}(G - e)$ is the maximum of the $\text{ecc}(v, G - e)$, it can be stored in $O(m)$ space with one entry for each edge e .

The second folklore FDO can only be used for undirected edge-weighted graphs. The FDO has stretch 2 and uses the fact that the diameter of the graph is intimately related to the eccentricity of *any* vertex. For an arbitrary v , we have that $\text{ecc}(v, G) \leq \text{diam}(G) \leq 2 \text{ecc}(v, G)$ as, by the triangle inequality, we can bound the distance between any two vertices $u, u' \in V$ by $d_G(u, u') \leq d_G(u, v) + d_G(v, u') \leq 2 \text{ecc}(v, G)$. The FDO again computes a shortest paths tree T rooted at a fixed source v and stores an array of length $n - 1$, corresponding to the edges of T . For every such edge e , one computes and stores $2 \text{ecc}(s, G - e)$. When queried with edge e , the FDO returns the stored value or, if e is not in the tree, the value $2 \text{ecc}(s, G)$. The size of this FDO is $O(n)$.

A maybe lesser-known way of building FDOs is via spanners. For any $\sigma > 0$, we say that a subgraph H of G is a *spanner* of stretch σ if, for every two vertices s, t of G , we have $d_H(s, t) \leq \sigma d_G(s, t)$. For every positive integer k , it is known how to construct a spanner H of G such that (a) H has a stretch of $2k - 1$ and (b) the size of H is $O(n^{1+1/k})$ [5]. Observe that for every edge $e = \{u, v\}$ that is in G but not in H , we have $d(u, v, e) \leq 2k - 1$. We now describe how spanners can be used to construct another easy oracle for undirected unweighted graphs whose stretch guarantee depends on both k and the inverse of $\text{diam}(G)$. This implies that the oracle already performs quite well for large-diameter graphs.

We construct such a spanner oracle with parameter k by first computing a spanner that satisfies (a) and (b). Then, we associate the value $\text{diam}(G - e)$ to each edge e in the spanner H and build a dictionary in which we store information about the edges of the spanner together with the corresponding associated values. Consider a query of edge e . If $e \in E(H)$ then we return the value associated with e ; otherwise, we return $\text{diam}(G) + 2(k - 1)$. The proof of the next lemma is deferred to the full version.

► **Lemma 8.** *For every positive integer k , the spanner oracle with parameter k has $O(n^{1+1/k})$ size, a constant query time, and a stretch of $1 + 2(k - 1)/\text{diam}(G)$.*

The result of Lemma 8 already implies the existence of sparse FDOs of $o(m)$ size and of stretch $\sigma < 3/2$ for sufficiently dense graphs with diameter strictly larger than 4. This does not contradict the lower bound of Theorem 1, but allows us to conclude that strong lower bounds on the size of FDOs for unweighted undirected graphs can only hold when the diameter of the input graph is bounded by a small constant.

3 Single Edge Failures

First, we treat single edge failures, $f = 1$. In this section, we assume the base graph G to be directed and present an $(1 + \varepsilon)$ -approximate fault-tolerant diameter oracle with space $O(m)$ and $O(1)$ query time. We give two variants, one is deterministic and combinatorial, the other randomized and algebraic. We then show that the space requirement is optimal up to the size of the machine word.

3.1 An $(1 + \varepsilon)$ -approximate FDO for Single Failures

We construct here the approximate FDO, thereby proving Theorem 2. Suppose we know for each $s, t \in V$ some shortest path $P(s, t)$ in G and additionally have access to a distance sensitivity oracle that, for any edge e , reports in constant time the replacement distances

$d(s, t, e)$ whenever needed. Clearly, $d(s, t, e)$ differs from the original graph distance only if e is on $P(s, t)$. To determine the diameters of *all* the graphs $G - e$, it is thus enough to query the DSO only for the edges on the shortest paths, which can be done in time $O(n^2 \cdot \text{diam}(G))$. We use approximation to avoid the cubic running time in case of a large diameter. For this, we randomly sample a small set B of so-called *pivots* and prove that it is enough to compute the replacement distances only between pairs from $B \times V$, instead of all pairs of vertices. Subsequently, we derandomize the pivot selection.

We fill in the details starting with the APSP computation in G and the preprocessing of the DSO. The combinatorial version uses a breath-first search from every vertex and the DSO of Bernstein and Karger [8], taking total time $\tilde{O}(mn)$. Alternatively, compute APSP algebraically and use the randomized DSO by Gu and Ren [28].⁴ APSP is computable in time $O(n^{2.575})$ on unweighted directed graphs with a variant of Zwick's algorithm [44, Corollary 4.5], this is in turn dominated by the $O(n^{2.5794})$ preprocessing time of the DSO [28]. After these computation, the distances $d(s, t)$, shortest paths $P(s, t)$ in G , and the replacement distances $d(s, t, e)$ are available to us (w.h.p., in the randomized case) with a constant query time per distance/path edge.

From here on out, the process for both variants is the same. Our fault-tolerant diameter oracle also allows non-edges to be queried, for which we return the original diameter $\text{diam}(G)$. To account for this, we store all edges in a static dictionary of size $O(m)$ that allows for worst-case constant look-up times after an $\tilde{O}(m)$ preprocessing [4, 29].⁵

Now fix a parameter $\varepsilon > 0$ for the approximation, possibly even depending on m, n . We initialize an array D indexed by the edges of G , all its cells hold the value $\text{diam}(G)$. Assume first that $\varepsilon \cdot \text{diam}(G) = O(\log n)$. For any two vertices s, t and edge e on the shortest path $P(s, t)$, we update $D[e]$ to the maximum of the previous value and $d(s, t, e)$. This takes $O(n^2 \text{diam}(G)) = \tilde{O}(n^2/\varepsilon)$ time. After all updates, the entry $D[e]$ stores the *exact* fault-tolerant diameter $\text{diam}(G - e)$ (possibly w.h.p.). For $\varepsilon \cdot \text{diam}(G) = \omega(\log n)$, we first give a randomized $(1+\varepsilon)$ -approximation and later derandomize it in Section 3.2. This yields the deterministic combinatorial algorithm of Theorem 3. The remaining use of randomness in the algebraic variant is due to the DSO by Gu and Ren [28].

To guard for the case that the failure of e disconnects the graph, we compute all *strong bridges* of G , that is, edges whose removal increases the number of strongly connected components, in time $O(m)$ with the algorithm by Italiano, Laura, and Santaroni [32]. For each strong bridge e , we set $D[e] = \infty$. To compute the other entries, we construct the set $B \subseteq V$ of pivots by randomly sampling every vertex independently with probability $C(\log n)/(\varepsilon \text{diam}(G))$ for a sufficiently large constant $C > 0$. A simple calculation using Chernoff bounds shows that $|B| = \tilde{O}(n/(\varepsilon \text{diam}(G)))$ w.h.p. Moreover, with high probability for all $s, t \in V$ and $e \in E$ such that $\varepsilon \text{diam}(G) < d(s, t, e) < \infty$, there exists a replacement path from s to t that avoids e and additionally contains a pivot from B . See [27, 41] for details. We update the entries of D in the same fashion as above, but now only use the (directed) distance $d(x, t, e)$ for all pivots $x \in B$ and vertices $t \in V$. In the end, we add $\varepsilon \text{diam}(G)$ to the value in $D[e]$. The array D is computable in time $O(m + n|B| \text{diam}(G)) = \tilde{O}(n^2/\varepsilon)$.

⁴ The DSO by Gu and Ren [28] is not path-reporting; if it were, we would not have to compute APSP. The fastest path-reporting algebraic DSO was given by Ren [38, 39] and can be constructed in time $O(n^{2.7233})$ on directed graphs, respectively in time $O(n^{2.6865})$ on undirected graphs.

⁵ The weak non-uniformity mentioned in [29], i.e., the need of compile-time constants depending on the word size, only holds if this size is $\omega(\log n)$, which is not the case for us.

We verify that $D[e]$ is an $(1+\varepsilon)$ -approximation of the fault-tolerant diameter $\text{diam}(G - e)$.

► **Lemma 9.** *We have $\text{diam}(G - e) \leq D[e] \leq (1+\varepsilon) \text{diam}(G - e)$ w.h.p.*

Proof. We can assume that $G - e$ is strongly connected as otherwise $D[e] = \infty = \text{diam}(G - e)$. The upper bound follows from $D[e] = \max_{x \in B, t \in V} d(x, t, e) + \varepsilon \text{diam}(G) \leq (1+\varepsilon) \text{diam}(G - e)$.

The main part consists of showing the lower bound $D[e] \geq \text{diam}(G - e)$. The idea is to prove the existence of a pivot $x \in B$ and vertex $t \in V$ whose replacement distance underestimates the fault-tolerant diameter by at most an additive term $\varepsilon \text{diam}(G)$, which we offset when computing $D[e]$. If $\text{diam}(G - e) \leq \varepsilon \text{diam}(G)$ (which can only happen for $\varepsilon \geq 1$), the lower bound holds vacuously as we have $D[e] \geq \varepsilon \text{diam}(G)$.

Let thus vertices $s, t \in V$ be such that $d(s, t, e) = \text{diam}(G - e) > \varepsilon \text{diam}(G)$. Since $G - e$ is strongly connected the diameter is finite and realized by some replacement path $P(s, t, e)$. In particular, we have $|P(s, t, e)| > \varepsilon \text{diam}(G)$. Let y be the unique vertex on $P(s, t, e)$ with $d(s, y, e) = \varepsilon \text{diam}(G)$. Recall that w.h.p. the set B hits *some* shortest path P' from s to y that avoids e . The path P' is not necessarily equal to the subpath $P(s, t, e)[s..y]$, but they have the same length $d(s, y, e)$. Substituting P' for $P(s, t, e)[s..y]$ therefore guarantees a replacement path from s to t that (w.h.p.) has a pivot $x \in B$ on its prefix of length $\varepsilon \text{diam}(G)$. For notational convenience, we use $P(s, t, e)$ to also denote this particular path.

The replacement distance from pivot x to target t satisfies $d(x, t, e) = |P(s, t, e)[x..t]| = |P(s, t, e)| - |P(s, t, e)[s..x]| \geq d(s, t, e) - \varepsilon \text{diam}(G)$. The entry $D[e]$ is also updated using the pivot x , whence $D[e] \geq d(x, t, e) + \varepsilon \text{diam}(G) \geq d(s, t, e) = \text{diam}(G - e)$. ◀

3.2 Derandomization

For the randomized combinatorial FDO, we had a preprocessing time of $\tilde{O}(mn + n^2/\varepsilon)$. The underlying APSP computation and the DSO are deterministic. We now derandomize the approximation part in the same asymptotic running time, proving Theorem 3. In Lemma 9, we used that the set B intersects at least one long replacement path from s to t exactly. We argue that it is in fact enough to hit the set of all vertices with distance at most $\varepsilon \text{diam}(G)$ from s in each strongly connected $G - e$. The pivot x does not need to be on any replacement path. The only assertion of Lemma 9 that is possibly in doubt is the lower bound $D[e] \geq \text{diam}(G - e)$. Let again s and t be such that $d(s, t, e) = \text{diam}(G - e)$ and let $x \in B$ be a pivot with $d(s, x, e) = d_{G - e}(s, x) \leq \varepsilon \text{diam}(G)$. Whenever $G - e$ is strongly connected, a replacement path $P(x, t, e)$ exists and, by the triangle inequality, we have $d(x, t, e) \geq d(s, t, e) - d(s, x, e) \geq d(s, t, e) - \varepsilon \text{diam}(G)$. The claim follows.

For the derandomization, we adopt the framework of Alon, Chechik and Cohen [3]. This involves efficiently finding a small set of critical paths such that hitting them ensures to hit each $(\varepsilon \text{diam}(G))$ -ball in the strongly connected $G - e$. If the critical paths are both short enough and few in numbers, it is then enough to compute the hitting set via the folklore greedy algorithm. In [3], it was sufficient to give a single set of critical paths. We generalize this to multiple sets, where the later-defined sets depend on the paths in the former.

Set $\ell = \min\{\varepsilon \text{diam}(G), \sqrt{n}\}$ and let r be an arbitrary vertex in G . We compute the in-tree $T_{\text{in}}(r)$, containing the shortest paths in G leading to r , with breath-first search. In the set \mathcal{P} , we collect, for each vertex s with $d(s, r) > \ell$, the path of $T_{\text{in}}(r)$ starting in s and having length ℓ . Let $P \in \mathcal{P}$ be a path with start vertex s and let $e \in E(P)$ be such that it is not a strong bridge. We compute the in-tree $T_{\text{in},e}(r)$ in $G - e$ rooted in r . Note that s has distance $d(s, r, e) \geq d(s, r) > \ell$ from the root in the tree. We add the corresponding path to the set \mathcal{P}_e . The original in-tree $T_{\text{in}}(r)$ contains only $n - 1$ edges, so all

trees can be computed in total time⁶ $O(mn)$. Moreover, there are at most $\ell + 1$ paths with starting vertex s . In total, we thus have $O(n\ell)$ paths each of length ℓ . A greedy algorithm computes a hitting set B for all paths in the \mathcal{P} and \mathcal{P}_e . It iteratively selects the vertex that is contained in the most yet unhit paths, it terminates in time $\tilde{O}(n\ell^2) = \tilde{O}(n^2)$ and produces a set of $|B| = \tilde{O}(n/\ell) = \tilde{O}(n/(\varepsilon \text{diam}(G)))$ pivots, see [3, 34]. We used the definition $\ell = \min\{\varepsilon \text{diam}(G), \sqrt{n}\}$ for both estimates. Finally, we add the root r to the set B to cover all paths in the trees that are shorter than ℓ .

► **Lemma 10.** *For each vertex $s \in V$ and edge e such that $G - e$ is strongly connected, there exists a pivot $x \in B$ with $d(s, x, e) \leq \varepsilon \text{diam}(G)$.*

Proof. If $d(s, r) \leq \varepsilon \text{diam}(G)$, we are done. Otherwise, let P be the prefix of length ℓ of the path from s to r in the tree $T_{\text{in}}(r)$, whence $P \in \mathcal{P}$. If P does not contain the edge e , it also exists in $G - e$ and the corresponding pivot $x \in B \cap V(P)$ satisfies $d(s, x, e) = d(s, x) \leq \ell \leq \varepsilon \text{diam}(G)$. If P contains e , then let instead $P' \in \mathcal{P}_e$ be the length- ℓ prefix of the path from s to r in $T_{\text{in},e}(r)$. Again, $x \in B \cap V(P')$ implies $d(s, x, e) \leq \varepsilon \text{diam}(G)$. ◀

3.3 Space Lower Bounds

Finally, we prove Theorem 1 thus showing that the space requirement of the FDOs in Theorems 2 and 3 is near-optimal provided that the stretch is $\sigma = \sigma(m, n) < 3/2$, that is, $\varepsilon < 1/2$. This even holds for the simpler task of computing the diameter in undirected graphs. For better exposition, we first show that any diameter oracle with such a stretch requires $\Omega(n^2)$ space on at least one n -vertex graph, which is, however, only tight for dense graphs. We then sparsify the construction to for an $\Omega(m)$ bound for graphs with m edges. Any σ -approximate FDO solves the promise problem of distinguishing, for each edge e , whether $G - e$ has diameter 2 or 3.

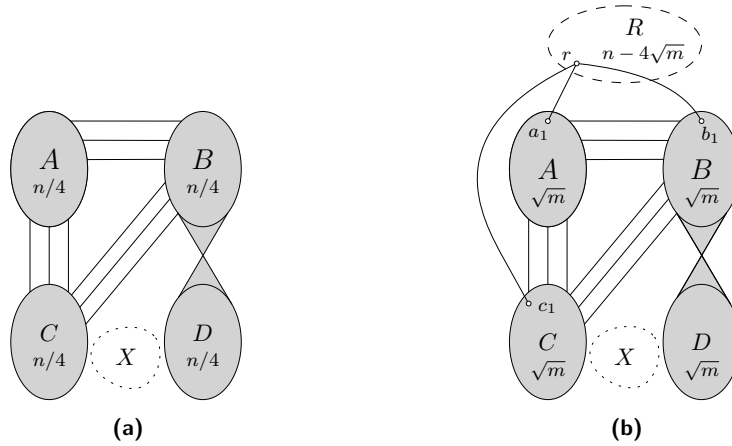
► **Lemma 11.** *There is a graph G on n vertices such that $G - e$ has diameter 2 or 3 for any $e \in E$. Any data structure that decides which one is the case must take $\Omega(n^2)$ bits of space.*

Proof. We give an incompressibility argument by encoding any binary $(n/4) \times (n/4)$ matrix X in the fault-tolerant diameters of G . No data structure can store this in $o(n^2)$ bits. The construction is illustrated in Figure 1a.

Without losing generality, n is divisible by 4, we can add up to three dummy vertices if needed. Split the vertex set equally into four groups A, B, C, D and let $a_1, \dots, a_{n/4}$ be an arbitrary numbering of the elements of A , same with the other groups. All groups are made into cliques and, for all $i \in [n/4]$, we make a_i, b_i , and c_i into a triangle. This results in matchings for the pairs (A, B) , (B, C) , and (A, C) , respectively. We further add edges so as to make (B, D) into a biclique. To encode the matrix X , we introduce the edge $\{c_i, d_j\}$ if and only if $X_{i,j} = 1$.

The graph G indeed has diameter 2 (even if X is the all-zeros matrix). Vertices a_i and b_j are joined by the path (a_i, a_j, b_j) —which by symmetry also holds for the other pairs of groups among A, B , or C —and the vertices a_i or c_i are connected to d_j via the paths (a_i, b_i, d_j) or (c_i, b_i, d_j) , respectively. Removing any edge increases the diameter by at most 1 since for any $e = \{u, v\}$ there exists a common neighbor in $w \in N[u] \cap N[v]$. This is clear inside the (bi-)cliques. For the matching edges, say $e = \{a_i, b_i\}$, we have $w = a_j, j \neq i$. Finally, for $e = \{c_i, d_j\}$ (if it exists), we have $w = b_i$.

⁶ For a single source, there are randomized algorithms known that compute the trees faster [11, 15, 26].



■ **Figure 1** Illustration of Lemma 11 **(a)** and of Lemma 12 **(b)**. The full ellipses A, B, C, D are cliques on the respective number of vertices, the dashed ellipse R is an independent set. The three parallel lines stand for matchings, the two crossed lines for a biclique. Edges encoding the binary matrix X run between C and D . Every vertex $r \in R$ is connected to $a_1 \in A, b_1 \in B$, and $c_1 \in C$.

We now prove that the graph $G - \{b_i, d_j\}$ has diameter 3 if and only if the edge $\{c_i, d_j\}$ is *not* present in G , that is, iff $X_{i,j} = 0$. When arguing the diameter above, edge $\{b_i, d_j\}$ was only needed for the paths (a_i, b_i, d_j) and (c_i, b_i, d_j) . Consider the neighborhoods of the three vertices in $G - \{b_i, d_j\}$, $N[a_i] = A \cup \{c_i, d_i\}$, $N[c_i] = C \cup \{a_i, b_i\} \cup \{d_k \mid X_{i,k} = 1\}$, and $N[d_j] = D \cup (B \setminus \{b_i\}) \cup \{c_k \mid X_{k,j} = 1\}$. If $X_{i,j} = 1$, then the neighborhoods intersect, namely in c_i , keeping the diameter at 2. If, however, $X_{i,j} = 0$, then $N[a_i] \cap N[d_j] = \emptyset$ and the diameter increases to 3. ◀

We now refine the result to give a better bound for sparse graphs. Note that a logarithmic gap remains between Lemma 12 and Theorem 2 since we lower bound the space at $\Omega(m)$ bits while the FDO takes this many words.

► **Lemma 12.** *There is a graph G with m edges such that $G - e$ has diameter 2 or 3 for any edge $e \in E$. A data structure that decides which one is the case must take $\Omega(m)$ bits of space.*

Proof. The main weakness of the construction in Lemma 11 is that it requires $\Omega(n^2)$ edges inside the cliques. As it turns out, this is not necessary and we can sparsify the graph G as long as we keep its diameter at 2. Figure 1b shows the idea of the sparsification.

Let m' be a parameter to be fixed later. We now store a binary $\sqrt{m'} \times \sqrt{m'}$ matrix X . Split the vertices into five groups, where A, B, C, D each contain $\sqrt{m'}$ vertices and R the remaining $n - 4\sqrt{m'}$. The edges among vertices in A through D are the same as in Lemma 11. Each vertex in R has degree 3 and is connected to a_1, b_1 , and c_1 . The graph G has $4\binom{\sqrt{m'}}{2} + 3\sqrt{m'} + m' + |\{(i, j) \mid X_{i,j} = 1\}| + 3(n - 4\sqrt{m'}) = O(m')$ edges. We fix the parameter m' such that the total number of edges is m . If needed, we introduce additional edges among vertices in R without affecting the result.

Note that the eccentricity of any vertex in $r \in R$ is 2 (even if R is not an independent set). Vertex a_i is reached via the path (r, a_1, a_i) , similar for the vertices in B and C , the ones in D are reached via b_1 . Moreover, for any edge involving r , say $\{r, a_1\}$, we have $b_1 \in N[r] \cap N[a_1]$. Therefore, the proof that G has diameter 2, $G - e$ has diameter 2 or 3, and $G - \{b_i, d_j\}$ has diameter 3 iff $X_{i,j} = 0$ is almost exactly as in Lemma 11. The sole difference is the case in which the edge $\{b_1, d_j\}$ fails since this may also increase the eccentricity of r . This is settled by observing that the neighborhood $N[r] = \{r, a_1, b_1, c_1\}$ in $G - \{b_1, d_j\}$ intersects $N[d_j]$ iff $X_{1,j} = 1$. To accommodate all possible matrices X , we require $\Omega(m') = \Omega(m)$ bits. ◀

The same construction shows that for edge-weighted graphs there is no $(2-\varepsilon)$ -approximate FDO, for any $\varepsilon = \varepsilon(m) > 0$, with space $o(m)$. In more detail, we choose an $\varepsilon' > 0$ small enough so that $\varepsilon' < 2\varepsilon/(1-\varepsilon)$ and give weight ε' to all matching edges as well as the edges incident to vertices in R , all other edges are weighted 2. One can verify that $\text{diam}(G) = 2 + \varepsilon'$ and the fault-tolerant diameter $\text{diam}(G - \{b_i, d_j\})$ remains at that value iff $\{c_i, d_j\}$ is present, it raises to $4 + \varepsilon'$ otherwise. The bound on the stretch cannot be improved as shown by the trivial FDO discussed in the introduction, which gives a 2-approximation in $O(n)$ space.

4 Multiple Edge Failures

We now turn to multiple edge failures. Recall that in the fault-tolerant setting the maximum number f of failures is known in advance, and stretch, space, preprocessing, and query time usually depend on f . In this section, we first prove the following lemma. Let $\alpha = \alpha(m, n)$ denote the inverse Ackermann function.

► **Lemma 13** (Theorem 5 with explicit logarithmic factors). *For every undirected graph with non-negative edge weights, there exists a deterministic combinatorial $(f+2)$ -approximate f -FDO that takes $O(fn \log^2 n)$ space and has $O(fm\alpha + fn \log^3 n)$ preprocessing time and $O(f^2 \log^2 n)$ query time. For $f = 1$, the size of the oracle is $O(n)$, the preprocessing time $O(m\alpha + n \log n)$, and the query time is constant.*

Bilò et al. [9] designed an $(2f+1)$ -approximate single-source f -DSO. That means, the oracle processes an undirected graph G with non-negative edge weights and a distinguished source s , and, upon query (t, F) with $|F| \leq f$, it returns $d(s, t, F)$. The oracle can be built in $O(fm\alpha + fn \log^3 n)$ time, has size $O(fn \log^2 n)$, and answers queries in $O(f^2 \log^2 n)$ time. In principle we can modify the oracle so as, when queried with the set F , it returns twice the eccentricity of s in the graph $G - F$. This would clearly allow us to construct an f -FDO of stretch $2 \cdot (2f+1)$. We show that the same oracle construction, but with a better query algorithm, allows us to develop an f -FDO of stretch $f+2$.

We let $w(e)$ denote the weight of the edge $e \in E$. The length of a path is now defined as the sum of its edge weights; the definitions of distance and diameter are adjusted accordingly. The oracle in [9] first computes a shortest path tree T of G rooted at the source s and uses it to re-weight all the edges of G . The new weight function w' assigns weight of 0 to each edge of T and weight $w'(e) = d(s, x) + w(e) + d(y, s)$ to any other edge $e = \{x, y\}$. When queried with (t, F) , the oracle computes a spanning forest T_F of $G - F$ w.r.t. the new weight function w' in $O(f^2 \log^2 n)$ time. Let $k = |F \cap E(T)|$. The oracle replaces the k failing edges in $F \cap E(T)$ with a minimum-weight set of edges in $G - F$ w.r.t. to w' , say E_F , whose addition to $T - F$ forms a spanning forest of $G - F$.⁷ The obtained forest T_F is then used to estimate the distance from s to t in $G - F$. We reuse a nice property proven in [9].

► **Lemma 14** (Bilò et al. [9]). *T_F is a minimum spanning forest of $G - F$ w.r.t. w' .*

Our query algorithm works as follows. Let tree T be rooted at s and $F \cap E(T) = \{f_1, \dots, f_k\}$ with $k \leq f$ the edges in T that are also in F . Let T_0, \dots, T_k denote the $k+1$ subtrees of $T - F$, and r_i the root of the subtree T_i . W.l.o.g., we assume $r_0 = s$. We use

⁷ This is done by computing, for each unordered pair $\phi = (T', T'')$ of connected components of $T - F$, the minimum-weight edge w.r.t. w' , say e_ϕ , that has one endpoint in T' and the other endpoint in T'' . Then, the set E_F is computed in $O(f^2)$ time using any time-efficient algorithm for computing a minimum spanning tree of an auxiliary graph in which each of the connected components of $T - F$ is modelled by a vertex and the edge between the unordered pair $\phi = (T', T'')$ of $T - F$ has a weight equal to $w'(e_\phi)$. The authors of [9] design a data structure that is able to retrieve, for each pair $\phi = (T', T'')$ of connected components of $T - F$, the edge e_ϕ in $O(\log^2 n)$ time.

f_1, \dots, f_k to compute the roots r_1, \dots, r_k in $O(k)$ time. We then build a forest T' on $k+1$ new vertices v_0, \dots, v_k , where v_i represents T_i . The forest T' contains an edge $\{v_i, v_j\}$ iff E_F contains an edge e with one end point in $V(T_i)$ and the other in $V(T_j)$. Obviously, if T' is not connected, then we can simply certify that $\text{diam}(G-F) = \infty$. So, we assume that T' is a tree. We root T' at v_0 and denote by e_i the edge that joins v_i with its parent $p(v_i)$. We compute the value $\Delta = \max_{1 \leq i \leq k} w'(e_i) - d(s, r_i)$ and output $\widehat{D} = f\Delta + 2 \cdot \max_{t \in V} d(s, t)$. The time needed for the query algorithm is dominated by the computation of T in time $O(f^2 \log^2 n)$ as all the new operations can be performed in $O(f^2)$ time. Observe that $\max_{t \in V} d(s, t)$ is independent of F and can be precomputed in time $O(n)$.

For a single failure, $f = 1$, the query time can be reduced to $O(1)$. In fact, for each edge e of T , it is enough to precompute the minimum weight edge of $E(G) \setminus E(T)$, w.r.t. weight function w' , that crosses the cut induced by $T - e$. This, a.k.a. the sensitivity analysis problem of a minimum spanning tree, can be solved in $O(m \log \alpha)$ time on a graph with m edges [37]. We show in the remainder that \widehat{D} is an $(f+2)$ -approximation of $\text{diam}(G-F)$. The proof of the following lemma can be found in the full version.

► **Lemma 15.** *We have that $\text{diam}(G-F) \geq \Delta$.*

We now prove the approximation with the help of Lemma 15.

► **Lemma 16.** *The value \widehat{D} satisfies $\text{diam}(G-F) \leq \widehat{D} \leq (f+2) \text{diam}(G-F)$.*

Proof. Again, we only need to prove anything if T' is connected, which implies that T_F is connected. By Lemma 15, we have that $\text{diam}(G-F) \geq \Delta$. Moreover, $\text{diam}(G-F) \geq \text{diam}(G) \geq \max_{t \in V} d(s, t)$. The value \widehat{D} returned by the query algorithm satisfies $\widehat{D} \leq f\Delta + 2 \max_{t \in V} d(s, t) \leq (f+2) \text{diam}(G-F)$. It remains to show that $\widehat{D} \geq \text{diam}(G-F)$. We prove the latter by verifying that, for any two vertices x and y , $\widehat{D} \geq d(x, y, F)$ holds.

Let r_x and r_y be the roots of the subtrees of $T-F$ that contain x and y , respectively. It is possible that $r_x = r_y$. Let $r_{p(i)}$ denote the root of the tree of $T-F$ that corresponds to the parent vertex $p(v_i)$ in T' . Consider the subgraph of T_F consisting of the edges of the paths in T_F between the following pairs of vertices: (a) $r_{p(i)}$ and r_i for every i , (b) x and r_x , (c) y and r_y . The subgraph contains a path from x to y since T_F is connected. Therefore, the replacement distance $d(x, y, F)$ is upper bounded by the total weight of the subgraph. The path in T_F between r_x and x has length at most $\max_{t \in V} d(s, t)$ as r_x is an ancestor of x in the shortest path tree T rooted at s ; same for r_y and y . Finally, for any $i > 0$, let $e_i = \{x_i, y_i\}$ be the edge in E_F that caused the addition of the edge $(v_i, p(v_i))$ in T' . W.l.o.g., we assume that x_i (resp., y_i) is a vertex of the tree of $T-F$ represented by v_i (resp., $p(v_i)$) in T' . The path from r_i to $r_{p(i)}$ in $G-F$ has length at most $d(r_i, x_i) + w(e_i) + d(y_i, r_{p(i)}) \leq d(r_i, x_i) + w(e_i) + d(y_i, s) + d(s, r_i) - d(s, r_i) = w'(e_i) - d(s, r_i) \leq \Delta$. Therefore, $d(x, y, F) \leq k\Delta + 2 \max_{t \in V} d(s, t) \leq f\Delta + 2 \max_{t \in V} d(s, t) = \widehat{D}$. ◀

4.1 Exact f -FDO for Low Diameter

We show that one can swap approximation for query time in low-diameter graphs, namely, with diameter at most $n^{\delta/f}/(f+1)$ for arbitrary $\delta = \delta(m, n) > 0$. This is summarized in Theorem 6. The case $f = 1$ is solved like in Section 3.1 only that there is no need for approximation here as the diameter is small enough to process all pairs of vertices in time $O(n^{2+\delta})$. We thus assume $f \geq 2$. We adapt a space-saving technique introduced by Chechik et al. [13]. In a bird's-eye view, we construct a recursion tree $T(s, t)$ of size $O(n^\delta)$ for each pair of vertices s and t . It contains all *relevant* replacement distances $d(s, t, F)$ for sets F with up to f failures. We then show how we can simulate the search for $\text{diam}(G-F)$ in the $O(n^2)$ trees in total time $O(2^f)$.

Afek et al. [1, Theorem 1] showed that if G is undirected, then any shortest path in $G - F$, with $|F|$, is a concatenation of at most $|F| + 1$ shortest paths in G . The condition on the diameter and $|F| \leq f$ ensure that every path below has length at most $(f+1) \cdot \text{diam}(G) \leq n^{\delta/f}$.

Assume we have access to a path-reporting f -DSO. That means, upon query (s, t, F) , the oracle either certifies that $d(s, t, F) = \infty$, i.e., s and t are disconnected in $G - F$, or reports the replacement distance and a shortest s - t -path in $G - F$. The preprocessing time of the combinatorial version is assumed to be $\tilde{O}(f m n^{1+\delta})$ with a $\tilde{O}(f n^{(1-1/f)\delta} + |P|) = \tilde{O}(f n^{(1-1/f)\delta})$ query time w.h.p. reporting path P . Here, we used the assumption $f \geq 2$, whence $|P| \leq n^{\delta/f} = \tilde{O}(f n^{(1-1/f)\delta})$. Alternatively, we have algebraic preprocessing in time $\tilde{O}(f n^{\omega+\delta})$. We show how to obtain the oracle in the full version, using an idea of Weimann and Yuster [43] with a more refined analysis of the query time.

Fix two vertices s and t . We construct the tree $T(s, t)$ recursively. Each node in the tree is associated with a set $F' \subseteq \binom{V}{2}$ containing $f' = |F'| \leq f$ possible failures. We have $F' = \emptyset$ in the root. Upon creation, the node queries the assumed oracle with (s, t, F') and holds the returned path $P(s, t, F')$, if any. If $f' = f$ or s and t are disconnected in $G - F'$, the node is a leaf. Otherwise, it has $d(s, t, F')$ many children, one for each edge of $e \in E(P(s, t, F'))$ of the path. The respective child is associated with the set $F' \cup \{e\}$.

The tree indeed has at least one node for every distinct replacement distance $d(s, t, F)$ with $|F| \leq f$. To see this, let F', F be two sets with $F' \subseteq F \subseteq \binom{V}{2}$. Clearly, we have $d(s, t, F') \leq d(s, t, F)$, but $d(s, t, F') < d(s, t, F)$ can only hold if $F \setminus F'$ contains an edge of the path $P(s, t, F')$ in the node associated with F' . The fan-out of each node is at most $n^{\delta/f}$, the height of the tree is f . For all $s, t \in V$, the trees thus have $O(n^{2+\delta})$ nodes in total and can be constructed with that many queries to the f -DSO in time $\tilde{O}(f n^{2+(2-1/f)\delta})$.

Consider the following naive algorithm to handle a query to the f -FDO for the fault-tolerant diameter $\text{diam}(G - F)$. Each tree $T(s, t)$ is searched individually starting in the root. The processing of a node depends on the associated set F' . If it is a leaf or the set $F \setminus F'$ is disjoint from the replacement path $P(s, t, F')$, then we return the length $d(s, t, F')$ of the path; otherwise, we recurse on all children associated with $F' \cup \{e\}$ for all edges $e \in (F \setminus F') \cap E(P(s, t, F'))$. By the argument as above, the maximum over all reported distances is indeed $\max_{s, t \in V; F' \subseteq F} d(s, t, F') = \text{diam}(G - F)$. This approach can be improved significantly by aggregating the values $\{d(s, t, F')\}_{s, t \in V}$ already at construction.

Observe that we never query the underlying f -DSO with a set F' that contains non-edges. We prepare a hash table H whose entries are indexed by subsets of E of size at most f . For every query (s, t, F') we compare the returned replacement distance with the value $H[F']$. If no such entry exists, we initialize it with $d(s, t, F')$; else, we update it to $\max\{H[F'], d(s, t, F')\}$. The final table has size $O(n^{2+\delta})$ and we discard the trees. The table H is constructible w.h.p. in time $O(n^{2+\delta})$, guaranteeing constant query time [20, 36]. However, to simulate the naive algorithm for the query F to the f -FDO, we have to check $H[F']$ for all $O(2^f)$ subsets $F' \subseteq F$ as we do not know which ones were used during construction.

4.2 Space Lower Bound

We conclude with the space lower bound of Theorem 4. It rules out any finite stretch in $o(fn)$ space for an arbitrary number f of failures. We use the fact that an f -FDO with finite stretch is able to decide whether the edges in F are a cut-set of the graph.

Assume for now that f is even. Let k be the largest integer such that $fk + 1 \leq n$. We construct a graph G as follows. It has vertices c, v_1, \dots, v_{fk} as well as $n - fk - 1$ auxiliary vertices. Define $E_i = \{\{v_i, v_j\} \mid 1 \leq |i - j| \leq f/2\}$. The edge set of G is $\bigcup_{i=1}^{fk} E_i$ together with all possible edges $\{c, u\}$, including to the auxiliaries. In other words, G consists of a star

centered at c with $n - 1$ leaves, and leaves v_i, v_j are joined by an edge iff their indices have difference at most $f/2$. Let set \mathcal{G} contain all spanning subgraphs of G that retain at least all star edges incident to c . Since $|E_i| = f$, there are $|\mathcal{G}| = 2^{(f-1)f/2} = 2^{\Omega(fn)}$ such subgraphs.

Let H be any subgraph in \mathcal{G} . For $i \neq j$ with $|i - j| \leq f/2$, define the set $F_{i,j} = (E_i \setminus \{v_i, v_j\}) \cup \{c, v_i\}$. Note that $F_{i,j}$ may contain non-edges. We have $|F_{i,j}| = f$ and evidently $\{v_i, v_j\}$ is present in H iff $H - F_{i,j}$ is connected. Any two f -FDOs for graphs in \mathcal{G} thus differ in at least one bit. For odd values $f \geq 3$, we emulate this using $f - 1$ failures.

For the remaining case $f = 1$, we use a different construction. W.l.o.g., n is even, connecting a single excess vertex to some other vertex in the graph is immaterial. The graph G contains two parallel paths P_1 and P_2 , each on $n/2$ vertices, respectively numbered from 1 to $n/2$. The graph also contains a matching M in which the i -th vertex of P_1 is matched with the i -th vertex of P_2 . Let \mathcal{G} be the set of all spanning subgraphs that have at least all the edges of P_1 and M . We have $|\mathcal{G}| = 2^{(n/2)-1} = 2^{\Omega(n)}$. Let $H \in \mathcal{G}$ and define e_i , with $i < n/2$, be the edge of P_1 between the i -th and $(i+1)$ -th vertices. The corresponding edge of P_2 is present in H if and only if $H - e_i$ is connected.

References

- 1 Yehuda Afek, Anat Bremler-Barr, Haim Kaplan, Edith Cohen, and Michael Merritt. Restoration by Path Concatenation: Fast Recovery of MPLS Paths. *Distributed Computing*, 15:273–283, 2002. doi:10.1007/s00446-002-0080-6.
- 2 Josh Alman and Virginia Vassilevska Williams. A Refined Laser Method and Faster Matrix Multiplication. In *Proceedings of the 32nd Symposium on Discrete Algorithms (SODA)*, pages 522–539, 2021. doi:10.1137/1.9781611976465.32.
- 3 Noga Alon, Shiri Chechik, and Sarel Cohen. Deterministic Combinatorial Replacement Paths and Distance Sensitivity Oracles. In *Proceedings of the 46th International Colloquium on Automata, Languages, and Programming, (ICALP)*, pages 12:1–12:14, 2019. doi:10.4230/LIPIcs.ICALP.2019.12.
- 4 Noga Alon and Moni Naor. Derandomization, Witnesses for Boolean Matrix Multiplication and Construction of Perfect Hash Functions. *Algorithmica*, 16:434–449, 1996. doi:10.1007/BF01940874.
- 5 Ingo Althöfer, Gautam Das, David P. Dobkin, Deborah Joseph, and José Soares. On Sparse Spanners of Weighted Graphs. *Discrete and Computational Geometry*, 9:81–100, 1993. doi:10.1007/BF02189308.
- 6 Bertie Ancona, Monika Henzinger, Liam Roditty, Virginia Vassilevska Williams, and Nicole Wein. Algorithms and Hardness for Diameter in Dynamic Graphs. In *Proceedings of the 46th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 13:1–13:14, 2019. doi:10.4230/LIPIcs.ICALP.2019.13.
- 7 Surender Baswana and Telikepalli Kavitha. Faster Algorithms for All-pairs Approximate Shortest Paths in Undirected Graphs. *SIAM Journal on Computing*, 39:2865–2896, 2010. doi:10.1137/080737174.
- 8 Aaron Bernstein and David R. Karger. A Nearly Optimal Oracle for Avoiding Failed Vertices and Edges. In *Proceedings of the 41st Symposium on Theory of Computing (STOC)*, pages 101–110, 2009. doi:10.1145/1536414.1536431.
- 9 Davide Bilò, Luciano Gualà, Stefano Leucci, and Guido Proietti. Multiple-Edge-Fault-Tolerant Approximate Shortest-Path Trees. In *Proceedings of the 33rd Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 18:1–18:14, 2016. doi:10.4230/LIPIcs.STACS.2016.18.
- 10 Édouard Bonnet. 4 vs 7 Sparse Undirected Unweighted Diameter is SETH-hard at Time $n^{4/3}$. In *Proceedings of 48th International Colloquium on Automata, Languages, and Programming, (ICALP)*, 2021. To appear.

- 11 Shiri Chechik and Sarel Cohen. Near Optimal Algorithms for the Single Source Replacement Paths Problem. In *Proceedings of the 30th Symposium on Discrete Algorithms (SODA)*, pages 2090–2109, 2019. doi:10.1137/1.9781611975482.126.
- 12 Shiri Chechik and Sarel Cohen. Distance Sensitivity Oracles with Subcubic Preprocessing Time and Fast Query Time. In *Proceedings of the 52nd Symposium on Theory of Computing (STOC)*, pages 1375–1388, 2020. doi:10.1145/3357713.3384253.
- 13 Shiri Chechik, Sarel Cohen, Amos Fiat, and Haim Kaplan. $(1 + \epsilon)$ -Approximate f -Sensitive Distance Oracles. In *Proceedings of the 28th Symposium on Discrete Algorithms (SODA)*, pages 1479–1496, 2017. doi:10.1137/1.9781611974782.96.
- 14 Shiri Chechik, Daniel H. Larkin, Liam Roditty, Grant Schoenebeck, Robert E. Tarjan, and Virginia Vassilevska Williams. Better Approximation Algorithms for the Graph Diameter. In *Proceedings of the 25th Symposium on Discrete Algorithms (SODA)*, pages 1041–1052, 2014. doi:10.1137/1.9781611973402.78.
- 15 Shiri Chechik and Ofer Magen. Near Optimal Algorithm for the Directed Single Source Replacement Paths Problem. In *Proceedings of the 47th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 81:1–81:17, 2020. doi:10.4230/LIPIcs.ICALP.2020.81.
- 16 Keerti Choudhary and Omer Gold. Extremal Distances in Directed Graphs: Tight Spanners and Near-Optimal Approximation Algorithms. In *Proceedings of the 31st Symposium on Discrete Algorithms (SODA)*, pages 495–514, 2020. doi:10.1137/1.9781611975994.30.
- 17 Fan Chung and Linyuan Lu. The Average Distances in Random Graphs with Given Expected Degrees. *Proceedings of the National Academy of Sciences*, 99:15879–15882, 2002. doi:10.1073/pnas.252631999.
- 18 Edith Cohen and Uri Zwick. All-Pairs Small-Stretch Paths. *Journal of Algorithms*, 38:335–353, 2001. doi:10.1006/jagm.2000.1117.
- 19 Camil Demetrescu, Mikkel Thorup, Rezaul Alam Chowdhury, and Vijaya Ramachandran. Oracles for Distances Avoiding a Failed Node or Link. *SIAM Journal on Computing*, 37:1299–1318, 2008. doi:10.1137/S0097539705429847.
- 20 Martin Dietzfelbinger, Anna R. Karlin, Kurt Mehlhorn, Friedhelm Meyer auf der Heide, Hans Rohnert, and Robert E. Tarjan. Dynamic Perfect Hashing: Upper and Lower Bounds. *SIAM Journal on Computing*, 23:738–761, 1994. doi:10.1137/S0097539791194094.
- 21 Ran Duan, Yong Gu, and Hanlin Ren. Approximate Distance Oracles Subject to Multiple Vertex Failures. In *Proceedings of the 32nd Symposium on Discrete Algorithms (SODA)*, pages 2497–2516, 2021. doi:10.1137/1.9781611976465.148.
- 22 Ran Duan and Seth Pettie. Dual-Failure Distance and Connectivity Oracles. In *Proceedings of the 20th Symposium on Discrete Algorithms (SODA)*, pages 506–515, 2009. URL: <https://dl.acm.org/citation.cfm?id=1496770.1496826>.
- 23 Ran Duan and Tianyi Zhang. Improved Distance Sensitivity Oracles via Tree Partitioning. In *Proceedings of the 15th Algorithms and Data Structures Symposium (WADS)*, pages 349–360, 2017. doi:10.1007/978-3-319-62127-2_30.
- 24 Tobias Friedrich and Anton Krohmer. On the Diameter of Hyperbolic Random Graphs. *SIAM Journal on Discrete Mathematics*, 32:1314–1334, 2018.
- 25 Pawel Gawrychowski, Haim Kaplan, Shay Mozes, Micha Sharir, and Oren Weimann. Voronoi Diagrams on Planar Graphs, and Computing the Diameter in Deterministic $\tilde{O}(n^{5/3})$ Time. In *Proceedings of the 29th Symposium on Discrete Algorithms (SODA)*, pages 495–514, 2018. doi:10.1137/1.9781611975031.33.
- 26 Fabrizio Grandoni and Virginia Vassilevska Williams. Improved Distance Sensitivity Oracles via Fast Single-Source Replacement Paths. In *Proceedings of the 53rd Symposium on Foundations of Computer Science (FOCS)*, pages 748–757, 2012. doi:10.1109/FOCS.2012.17.
- 27 Fabrizio Grandoni and Virginia Vassilevska Williams. Faster Replacement Paths and Distance Sensitivity Oracles. *ACM Transaction on Algorithms*, 16:15:1–15:25, 2020. doi:10.1145/3365835.

- 28 Yong Gu and Hanlin Ren. Constructing a Distance Sensitivity Oracle in $O(n^{2.5794}M)$ Time. In *Proceedings of the 48th International Colloquium on Automata, Languages, and Programming (ICALP)*, 2021. To appear.
- 29 Torben Hagerup, Peter Bro Miltersen, and Rasmus Pagh. Deterministic Dictionaries. *Journal of Algorithms*, 41:69–85, 2001. doi:10.1006/jagm.2001.1171.
- 30 Monika Henzinger, Andrea Lincoln, Stefan Neumann, and Virginia Vassilevska Williams. Conditional Hardness for Sensitivity Problems. In *Proceedings of the 8th Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 26:1–26:31, 2017. doi:10.4230/LIPIcs.ITCS.2017.26.
- 31 Remco van der Hofstad. *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, UK, 2016. doi:10.1017/9781316779422.
- 32 Giuseppe F. Italiano, Luigi Laura, and Federico Santaroni. Finding Strong Bridges and Strong Articulation Points in Linear Time. *Theoretical Computer Science*, 447:74–84, 2012. doi:10.1016/j.tcs.2011.11.011.
- 33 Telikepalli Kavitha. Faster Algorithms for All-Pairs Small Stretch Distances in Weighted Graphs. *Algorithmica*, 63:224–245, 2012. doi:10.1007/s00453-011-9529-y.
- 34 Valerie King. Fully Dynamic Algorithms for Maintaining All-Pairs Shortest Paths and Transitive Closure in Digraphs. In *Proceedings of the 40th Symposium on Foundations of Computer Science (FOCS)*, pages 81–91, 1999. doi:10.1109/SFFCS.1999.814580.
- 35 Jon M. Kleinberg. Navigation in a Small World. *Nature*, 406:845–845, 2000. doi:10.1038/35022643.
- 36 Rasmus Pagh and Flemming Friche Rodler. Cuckoo Hashing. *Journal of Algorithms*, 51:122–144, 2004. doi:10.1016/j.jalgor.2003.12.002.
- 37 Seth Pettie. Sensitivity Analysis of Minimum Spanning Trees in Sub-Inverse-Ackermann Time. *Journal of Graph Algorithms and Applications*, 19:375–391, 2015. doi:10.7155/jgaa.00365.
- 38 Hanlin Ren. Improved Distance Sensitivity Oracles with Subcubic Preprocessing Time. In *Proceedings of the 28th European Symposium on Algorithms (ESA)*, pages 79:1–79:13, 2020. doi:10.4230/LIPIcs.ESA.2020.79.
- 39 Hanlin Ren. Improved Distance Sensitivity Oracles with Subcubic Preprocessing Time. *CoRR*, abs/2007.11495, 2020. ArXiv preprint. Full version of [38]. arXiv:2007.11495.
- 40 Liam Roditty. Approximating the Diameter. In Ming-Yang Kao, editor, *Encyclopedia of Algorithms*, pages 116–117. Springer, New York City, NY, USA, 2016. doi:10.1007/978-1-4939-2864-4_566.
- 41 Liam Roditty and Uri Zwick. Replacement Paths and k Simple Shortest Paths in Unweighted Directed Graphs. *ACM Transaction on Algorithms*, 8:33:1–33:11, 2012. doi:10.1145/2344422.2344423.
- 42 Mikkel Thorup and Uri Zwick. Approximate Distance Oracles. In *Proceedings on 33rd Symposium on Theory of Computing (STOC)*, pages 183–192, 2001. doi:10.1145/380752.380798.
- 43 Oren Weimann and Raphael Yuster. Replacement Paths and Distance Sensitivity Oracles via Fast Matrix Multiplication. *ACM Transactions on Algorithms*, 9:14:1–14:13, 2013. doi:10.1145/2438645.2438646.
- 44 Uri Zwick. All Pairs Shortest Paths Using Bridging Sets and Rectangular Matrix Multiplication. *Journal of the ACM*, 49:289–317, 2002. doi:10.1145/567112.567114.