

Parallel Graph Decompositions Using Random Shifts *

Gary L. Miller
CMU
glmiller@cs.cmu.edu

Richard Peng †
CMU
yangp@cs.cmu.edu

Shen Chen Xu
CMU
shenchex@cs.cmu.edu

ABSTRACT

We show an improved parallel algorithm for decomposing an undirected unweighted graph into small diameter pieces with a small fraction of the edges in between. These decompositions form critical subroutines in a number of graph algorithms. Our algorithm builds upon the shifted shortest path approach introduced in [Blelloch, Gupta, Koutis, Miller, Peng, Tangwongsan, SPAA 2011]. By combining various stages of the previous algorithm, we obtain a significantly simpler algorithm with the same asymptotic guarantees as the best sequential algorithm.

Categories and Subject Descriptors

F.2 [Theory of Computation]: Analysis of Algorithms and Problem Complexity

Keywords

Parallel algorithms; Low-diameter decompositions; Graph partitioning

1. INTRODUCTION

Graph decomposition aims to partition the vertices of a graph into well connected pieces so that few edges are between pieces. A variety of measures of the connectivity within a piece, such as diameter, conductance, and spectral properties have been studied. The more intricate measures such as conductance have proven to be particularly useful in applications [25], and have been well studied [20, 24]. However, these algorithms, as well as many others, use simpler low diameter decompositions as a subroutine. This variant takes a much more simplistic view of the connectivity within each piece, and measures it using only the diameter.

*Partially supported by the National Science Foundation under grant number CCF-1018463 and CCF-1065106

†Supported by a Microsoft Fellowship

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SPAA'13, July 23–25, 2013, Montréal, Québec, Canada.

Copyright 2013 ACM 978-1-4503-1572-2/13/07 ...\$15.00.

The various shortcomings of such decompositions as a stand-alone routine are offset by its potential as an algorithmic tool. Low diameter decompositions/clustering were first introduced in [4, 6], and have been used as core subroutine for a number of algorithms such as: approximations to sparsest cut [20, 24]; construction of spanners [12]; parallel approximations of shortest path in undirected graphs [13]; and generating low-stretch embedding of graphs into trees [3, 16, 15, 2].

More recently, the connection of low diameter decomposition with generating low stretch spanning trees was used in [9] to give nearly-linear work parallel solvers for SDD linear systems. These solvers can in turn be used as a black-box in algorithms for computing maximum flow and negative-length shortest path [11, 14]. This led to parallel algorithms whose work is within polylog factors of the best known sequential algorithms. As these are problems for which work-efficient parallelizations have proven to be elusive, parallel solvers for SDD linear systems represent a promising new direction for designing parallel algorithms.

The parallel SDD linear system solver algorithm from [9] is of mostly theoretical interest due to a large polylog factor in work. Much of this is due to the nearly-linear work, parallel low diameter decomposition algorithm introduced in the same paper, which was in turn used as a subroutine to generate tree embeddings. Therefore, finding improved parallel graph decomposition routines represent a natural direction for finding faster parallel solver algorithms.

In order to formally specify the diameter of a piece, it is crucial to emphasize the distinction between weak and strong diameter. The diameter of a piece $S \subseteq V$ can be defined in two ways, weak and strong diameter. Both of them define diameter to the maximum length of a shortest path between two vertices in S , while the difference is in the set of allowed paths. **Strong diameter** restricts the shortest path between two vertices in S to only use vertices in S , while **weak diameter** allows for shortcuts through vertices in $V \setminus S$. The optimal tree metric embedding algorithm [16] relies on weak diameter. It has been parallelized with polylog work overhead [10], but takes quadratic work.

A trend in algorithms that use weak diameter is that their running time tends to be quadratic. This is also the case with parallel algorithms for computing low diameter decompositions [5]. To date, nearly-linear work algorithms for finding tree embedding use strong diameter instead. While this leads to more difficulties in bounding diameters, the overall work is easier to bound since each piece certifies its own diameter, and therefore does not need to examine other pieces.

For SDD linear system solvers [10], strong diameter is also crucial because the final tree (which the graph embeds into) is formed by combining the shortest path tree in each of the pieces. As a result, we will use diameter of a piece to denote **strong diameter** for the rest of this paper, and define a low diameter decomposition as follows:

DEFINITION 1.1. *Given an undirected, unweighted graph $G = (V, E)$, a (β, d) decomposition is a partition of V into subsets $S_1 \dots S_k$ such that:*

- *The (strong) diameter of each S_i is at most d .*
- *The number of edges with endpoints belonging to different pieces is at most βm .*

A standard choice of parameters for such decompositions is $(\beta, O(\frac{\log n}{\beta}))$, which are in some sense optimal. Furthermore, when computing tree embedding, β is often set to $\log^{-c} n$ for some constant c . As a result, the algorithm given in [9], as well as our algorithm are geared towards small diameters. This makes the running time of these algorithms more than the \mathcal{NC} algorithms such as [5] for large values of β . However, they suffice for the purpose of generating tree/subgraph embedding in polylog depth, as well as low work parallel graph algorithms.

Obtaining these decompositions in the sequential setting can be done via a process known as ball growing. This process starts with a single vertex, and repeatedly adds the neighbors of the current set into the set. It terminates when the number of edges on the boundary is less than a β fraction of the edges within, which is equivalent to the piece having small conductance. Once the first piece is found, the algorithm discards its vertices and repeats on the remaining graph. The final bound of βm edges between the pieces can in turn be obtained by summing this bound over all the pieces. Using a consumption argument, one could prove that the diameter of a piece does not exceed $O(\frac{\log n}{\beta})$. Because we are okay with a depth that depends on $1/\beta$, and the piece's diameters can be bounded by $O(\frac{\log n}{\beta})$, finding a single piece is easy to parallelize. However, the strong diameter requirement means that we cannot start finding the second ball until we are done finding the first. This leads to a chain of sequential dependencies that may be as long as $\Omega(n)$, and is the main challenge in obtaining a parallel decomposition algorithm.

The parallel SDD linear system solver algorithm given in [9] relied on a parallel algorithm that computes a $O(\beta, \frac{\log^4 n}{\beta})$ decomposition in $O(\frac{\log^3 n}{\beta})$ depth and $O(m \log^2 n)$ work. This algorithm showed that some of the ball growing steps can be performed simultaneously in parallel, leading to balls which have small overlap. Then a randomly shifted shortest path routine is used to resolve these overlaps. In this paper we show that these two steps can be combined in a simple, global routine. This leads to a simple algorithm that picks random shifts in a more intricate way and assigns vertices to pieces using one shortest path invocation. In the PRAM model, our result can be described by the following theorem:

THEOREM 1.2. *There is an algorithm PARTITION that takes an unweighted graph with n vertices, m edges, a parameter $\beta \leq 1/2$ and produces a $(\beta, O(\frac{\log n}{\beta}))$ decomposition in expected $O(\frac{\log^2 n}{\beta})$ depth and $O(m)$ work.*

We will give an overview of our algorithm in Section 2 and define our notations and give relevant background in Section 3. Section 4 contains the analysis of our partition routine, and modifications to make it more suitable for implementation and parallelization are given in Section 5. In Section 6 we discuss some possible extensions to our algorithm.

2. OVERVIEW AND RELATED WORKS

In this section we give an intuitive view of our partition routine and discuss how it relates to various other graph decomposition algorithms that have been studied in the past. We will defer the implementation details to Section 5. A simple interpretation of our algorithm executing in parallel is outlined in Algorithm 1.

Algorithm 1 Parallel Partition Algorithm

PARALLEL PARTITION

Input: Undirected, unweighted graph $G = (V, E)$, parameter $0 < \beta < 1$ and parameter d indicating failure probability.

Output: $(\beta, O(\log n/\beta))$ decomposition of G with probability at least $1 - n^{-d}$.

- 1: *IN PARALLEL* each vertex u picks δ_u independently from an exponential distribution with mean $1/\beta$.
 - 2: *IN PARALLEL* compute $\delta_{\max} = \max\{\delta_u \mid u \in V\}$
 - 3: Perform *PARALLEL BFS*, with vertex u starting when the vertex at the head of the queue has distance more than $\delta_{\max} - \delta_u$.
 - 4: *IN PARALLEL* Assign each vertex u to point of origin of the shortest path that reached it in the BFS.
-

Steps 1, 2, and 4 of our algorithm are done independently at each vertex and are clearly parallelizable. So the main algorithmic aspects of our algorithm is in step 3, which is performing a breadth first search while recording the point of origin. Such processes have been well-studied [21, 8], and we will discuss how to use such routines in a black-box manner in Section 5. More intuitively, our algorithm can also be viewed as performing parallel ball growing with random delays. Each vertex u picks a start time according to some distribution, and if u is not already part of some other cluster at that time, u starts a cluster of its own and performs a breadth first search. The search takes one unit of time to propagate across an edge, and each such time steps can be performed in parallel over all vertices. If a vertex v visited during the search is not yet part of any other cluster, it joins the cluster of the vertex that first reached it, and its neighbors are added to the BFS queue. The randomized start times leads to both required properties of the decompositions, as well as small bound on the depth of the BFS trees. Figure 1 shows the resulting partitions of our algorithm with a 1000×1000 square grid as input and different values of β used to generate the delays. Lower β leads to larger diameter and fewer edges on the boundaries, which matches our more detailed analysis in Section 4.

To our knowledge, low diameter decompositions as stated in Definition 1.1 were first used for distributed algorithms in [4]. Subsequently it has been used as a key subroutine in the construction of low-stretch spanning trees algorithms [3], or more generally embeddings of graphs into trees [7]. Another application of unweighted decompositions is for efficiently computing separators in minor-free graphs [23, 28].

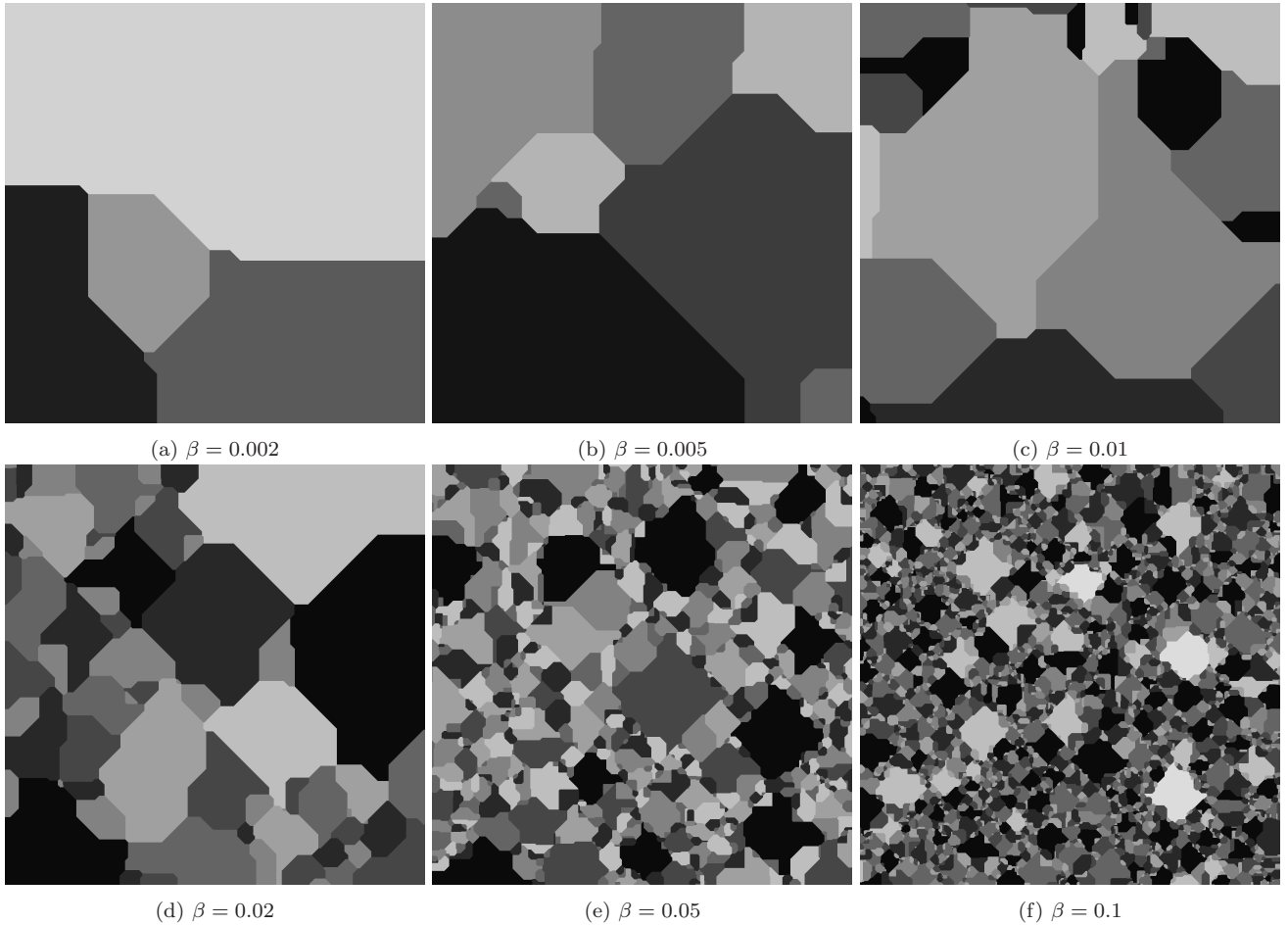


Figure 1: Decompositions generated by our algorithm on a 1000×1000 grid under varying values of β . Different colors represent different clusters

Our algorithm can be directly substituted into these algorithms, although the main improvements that we obtain are for generating low stretch spanning trees using the framework of [9].

A definition related to low diameter decomposition is block decompositions from [22]. One of their main algorithmic routines is to partition a graph into $O(\log n)$ blocks such that each connected piece in a block has diameter $O(\log n)$. This decomposition can also be obtained by iteratively running a $(\frac{1}{2}, O(\log n))$ low diameter decomposition $O(\log n)$ times. This is because the number of edges not in a block decreases by a factor of 2 per iteration.

The main scheme of our algorithm is to pick radii of the balls independently from some distribution. Similar approaches have been used in computing block decompositions in [22], as well as finding (r, ρ, γ) -probabilistic partitions needed for the Bartal trees [7]. Our partition scheme differs from these in that the process behaves identically on all vertices, and our guarantees are in terms of strong diameter. The first difference means that the formation of clusters in our algorithm does not have sequential dependencies; while the later actually leads to a lower work term. A typical method for meeting weak diameter requirements is to broadcast to all vertices within a certain distance [4, 22]. As the graph may have small diameter, this can lead to work

that is quadratic in the number of vertices. By broadcasting only along shortest paths in a way that is akin to breadth first search, we are able to reduce this to $O(m)$ work.

Aside from being directly related to decomposition schemes for unweighted undirected graphs, our algorithm can also be adapted to give guarantees similar to other improved decomposition schemes. When viewed as a sequential algorithm, it can also lead to similar guarantees on weighted graphs to the decomposition scheme from [7] as well as generalizations needed for improved low stretch spanning tree algorithms [7, 15]. However, the parallel performance of our algorithm in the weighted setting is less clear, and we will describe this question in more detail in Section 6.

3. BACKGROUND AND NOTATIONS

In this section we state some standard notations, and review some key ideas introduced in the Blelloch et al. algorithm [9]. Given a graph G , we use $\text{dist}(u, v)$ to denote the length of the shortest path from u to v . As with earlier works on tree embedding, we will pick a special vertex in each piece, and use the distance to the farthest vertex from it as an estimate for the diameter. This simplification can be made since the graph is undirected and the final bound allows for constant factors (specifically 2). We will denote

this special vertex the center of the piece, and denote the piece centered at u using S_u .

As the number of pieces in the final decomposition may be large (e.g. the line graph), a parallel algorithm needs to construct a number of pieces simultaneously. On the other hand, for closely connected graphs such as the complete graph, a single piece may contain the entire graph. As a result, if too many pieces are grown independently, the total work may become quadratic. The decomposition algorithm by Bleloch et al. [9] addressed this tradeoff by gradually increasing the number of centers picked iteratively. It was motivated by the (β, W) decompositions used in an algorithm by Cohen for approximating shortest paths in undirected graphs [13]. By running iterations with gradually more centers, it can be shown that the resulting pieces at each iteration have small overlap. This overlap is in turn resolved using a shifted shortest path algorithm, which introduces shifts (denoted by δ) at the centers and assigns vertex v to S_u that minimizes the shifted distance:

$$\text{dist}_{-\delta}(u, v) = \text{dist}(u, v) - \delta_u. \quad (1)$$

It was shown that by picking shifts uniformly from a sufficiently large range, a $(\beta, O(\frac{\log^c n}{\beta}))$ decomposition can be obtained.

Our algorithm can be viewed as a more streamlined algorithm that combines these two components. Note that sampling vertices with exponentially increasing density can be emulated by adding a large, step-like increase to the shifts of centers picked in earlier iterations. Furthermore, the need to have exponentially decreasing number of centers in the iterations suggests that the exponential distribution can be used in place of the (locally) uniform distribution. This distribution has been well-studied, and the properties of it that we will need have been used to study its order statistics in fault tolerance [27]. For a parameter γ , this distribution is defined by the density function:

$$f_{Exp}(x, \gamma) = \begin{cases} \gamma \exp(-\gamma x) & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We will denote it using $Exp(\gamma)$ and will also make use of its cumulative density function:

$$F_{Exp}(x, \gamma) = \Pr[Exp(\gamma) \leq x] = \begin{cases} 1 - \exp(-\gamma x) & \text{if } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

A crucial fact about the exponential distribution is that it is memoryless. That is, if we condition on $Exp(\gamma) \geq t$, then $Exp(\gamma) - t$ will follow the same distribution. We will also use order statistic of random independent variables following the exponential distribution. Given n random variables $X_1 \dots X_n$, the i^{th} order statistic of them is the value of the i^{th} smallest random variable. Another property of exponential distributions is that the difference between its order statistics also follow exponential distributions. The following fact as stated on page 19 of [17] has been used in a variety of settings.

FACT 3.1. *The n variables $X_{(1)}, X_{(2)} - X_{(1)}, \dots, X_{(n)} - X_{(n-1)}$ are independent and the density of $X_{(k+1)} - X_{(k)}$ is given by $f_{Exp}(x, \gamma)$ where $\gamma = (n - k)$.*

Let $X_{(i)}^n$ denote the i^{th} order statistic of n i.i.d. exponential random variables. An intuitive way to prove the above

fact is that by the i.i.d assumption, the cumulative density distribution of $X_{(1)}^n$ is given by

$$\begin{aligned} F_{X_{(1)}^n}(x) &= 1 - (1 - F(x))^n \\ &= 1 - \exp(-n\beta x) \end{aligned}$$

where $F(x) = 1 - \exp(-\beta x)$ is the cumulative distribution of $Exp(\beta)$. This shows that $X_{(1)}^n$ has an exponential distribution with mean $1/(n\beta)$. Conditioning on $X_{(1)}^n$, we get that $X_{(2)}^n - X_{(1)}^n$ is again an exponential distribution equal to $X_{(1)}^{n-1}$ because of the memoryless property of exponential distributions. We can repeat this argument to get the density of $X_{(k+1)}^n - X_{(k)}^n$ for all k up to $n - 1$ [17].

4. ANALYSIS

In this section, we will show that our partition routine indeed constructs a $(\beta, O(\frac{\log n}{\beta}))$ decomposition. For the purpose of this proof, we use a slightly different formulation of our algorithm given in Algorithm 2. In this view, our algorithm picks shifts δ_u for all vertices from independent exponential distributions with parameter β , and then assigns each vertex to a piece so that the shifted distances defined in (1) to the center of that piece is minimized.

Algorithm 2 Partition Algorithm Using Exponentially Shifted Shortest Paths

PARTITION

Input: Undirected, unweighted graph $G = (V, E)$, parameter β and parameter d indicating failure probability.

Output: $(\beta, O(\log n/\beta))$ decomposition of G with probability at least $1 - n^{-d}$.

- 1: For each vertex u , pick δ_u independently from $Exp(\beta)$
 - 2: Compute S_u by assigning each vertex v to the vertex that minimizes $\text{dist}_{-\delta}(u, v)$, breaking ties lexicographically
 - 3: **return** $\{S_u\}$
-

We start by showing that the assignment process readily leads to bounds on strong diameter. Specifically, the strong diameter of S_u can be measured using distances from u in the original graph.

LEMMA 4.1. *If $v \in S_u$ and v' is the last vertex on the shortest path from u to v , then $v' \in S_u$ as well.*

Proof The proof is by contradiction, suppose v' belongs to $S_{u'}$ for some $u' \neq u$. The fact that v' is the vertex before v on the shortest path from u implies $\text{dist}_{-\delta}(u, v) = \text{dist}_{-\delta}(u, v') + 1$. Also, as v' is adjacent to v , we also have $\text{dist}_{-\delta}(u', v) \leq \text{dist}_{-\delta}(u', v') + 1$. Since v' belongs to $S_{u'}$ instead of S_u , we must have one of the following two cases:

1. v' is strictly closer to u' than u in terms of shifted distance. In this case we have $\text{dist}_{-\delta}(u', v') < \text{dist}_{-\delta}(u, v')$, which when combined with the conditions above gives:

$$\begin{aligned} \text{dist}_{-\delta}(u', v) &\leq \text{dist}_{-\delta}(u', v') + 1 \\ &< \text{dist}_{-\delta}(u, v') + 1 \\ &= \text{dist}_{-\delta}(u, v). \end{aligned}$$

So v is strictly closer to u' than u as well, which implies that v should not be assigned to S_u .

2. The shifted distances are the same, and u' is lexicographically earlier than u . Here a similar calculation gives $\text{dist}_{-\delta}(u', v) \leq \text{dist}_{-\delta}(u, v)$. If the inequality holds strictly, we are back to the case above. In case of equality, the assumption that u' is lexicographically earlier than u means v should not be in S_u as well. ■

Note that the second case is a zero probability event, and its proof is included to account for roundings in implementations that we will describe in Section 5.

To bound the strong diameter of the pieces, it suffices to bound the distance from a vertex to the center of the piece that it is assigned to. Since any vertex $v \in S_u$ could have been potentially included in S_v , the shift value of the center δ_u serves as an upper bound on the distance to any vertex in S_u . Therefore, $\delta_{\max} = \max_u \delta_u$ serves as an upper bound for the diameter of each piece. Its expected value and concentration can be bounded as follows.

LEMMA 4.2. *The expected value of the maximum shift value is given by H_n/β where H_n is the n th harmonic number. Furthermore, with high probability, $\delta_u \leq O(\frac{\log n}{\beta})$ for all vertices u .*

Our proof below follows closely following the presentation in Chapter 1.6. of [17].

Proof The expected value can be found by summing over the differences in order statistics given in Fact 3.1.

$$\begin{aligned} \mathbf{E} \left[\max_{u \in V} \delta_u \right] &= \mathbf{E} [X_{(n)}^n] \\ &= \frac{1}{\beta} \sum_{i=1}^n \frac{1}{i} \\ &= \frac{H_n}{\beta}. \end{aligned}$$

For the concentration bound, by the cumulative distribution function of the exponential distribution the probability of $\delta_u \geq (d+1) \cdot \frac{\ln n}{\beta}$ is:

$$\begin{aligned} \exp \left(-(d+1) \cdot \beta \frac{\ln n}{\beta} \right) &= \exp(-(d+1) \ln n) \\ &\leq n^{-(d+1)}. \end{aligned}$$

Applying union bound over the n vertices then gives the bound. ■

The other property that we need to show is that few edges are between the pieces. We do so by bounding the probability of two endpoints of an edge being assigned to two different pieces. In order to keep symmetry in this argument, it is helpful to consider shifted distances from a vertex to the midpoint of an edge. This slight generalization can be formalized by replacing an edge uv with two length $1/2$ edges, uw and wv . We first show that an edge's end points can be in different pieces only if there are two different vertices whose shifted shortest path to its midpoint are within 1 of the minimum.

LEMMA 4.3. *Let uv be an edge with midpoint w such that when partitioned using shift values δ , $u \in S_{u'}$ and $v \in S_{v'}$. Then both $\text{dist}_{-\delta}(u', w)$ and $\text{dist}_{-\delta}(v', w)$ are within 1 of the minimum shifted distance to w .*

Proof Let the pieces that contain u and v be $S_{u'}$ and $S_{v'}$ respectively ($u' \neq v'$). Let the minimizer of $\text{dist}_{-\delta}(x, w)$ be w' . Since w is distance $1/2$ from both u and v , we have

$$\text{dist}_{-\delta}(w', u), \text{dist}_{-\delta}(w', v) \leq \text{dist}_{-\delta}(w', w) + 1/2.$$

Suppose $\text{dist}_{-\delta}(u', w) > \text{dist}_{-\delta}(w', w) + 1$, then we have:

$$\begin{aligned} \text{dist}_{-\delta}(u', u) &\geq \text{dist}_{-\delta}(u', w) - 1/2 \\ &> \text{dist}_{-\delta}(w', w) + 1/2 \\ &\geq \text{dist}_{-\delta}(w', u), \end{aligned}$$

a contradiction with u' being the minimizer of $\text{dist}_{-\delta}(x, u)$. The case with v follows similarly. ■

An even more accurate characterization of this situation can be obtained using the additional constraint that the shortest path from w' to w must go through one of u or v . However, this lemma suffices for abstracting the situation further to applying random decreases $\delta_1, \delta_2 \dots \delta_n$ to a set of numbers $d_1 \dots d_n$ corresponding to $\text{dist}(x, w)$. We now turn our attention to analyzing the probability of another shifted value being close to the minimum when shifts are picked from the exponential distribution.

The memoryless property of the exponential distribution gives an intuitive way to bound this probability. Instead of considering the vertices picking their shift values independently, consider them as light bulbs with lifetime distributed according to $\text{Exp}(\beta)$, and the d_i s indicate the time each light bulb is being turned on. Then $\min_i d_i - \delta_i$ corresponds to the time when the last light bulb burns out, and we want to bound the time between that and the second last. In this setting, the memoryless property of exponentials gives that when the second to last light bulb fails, the behavior of the last light bulb does not change and its lifetime after that point still follows the same distribution. Therefore, the probability that the difference between these two is less than c can be bounded using the cumulative distribution function:

$$\begin{aligned} 1 - \exp(-c\beta) &\approx 1 - (1 - c\beta) \quad (\text{When } c\beta \text{ is small}) \\ &= c\beta. \end{aligned}$$

The only case that is not covered here is when the last light bulb has not been turned on yet when the second last failed. However, in that case this probability can only be less. We give a rigorous version of this intuitive proof below. An algebraic proof using the definition of the exponential distribution can be found in Appendix A.

LEMMA 4.4. *Let $d_1 \leq \dots \leq d_n$ be arbitrary values and $\delta_1 \dots \delta_n$ be independent random variables picked from $\text{Exp}(\beta)$. Then the probability that the smallest and the second smallest values of $d_i - \delta_i$ are within c of each other is at most $O(\beta c)$.*

Proof It is more convenient to consider the differences between the largest and second largest of the negations of the shifted values, $-(d_i - \delta_i)$. Let d'_i denote $-d_i$, by the assumption of $d_1 \leq \dots \leq d_n$ we have $d'_1 \geq \dots \geq d'_n$. Define $X_i = d'_i + \delta_i - d'_1$ and let $X_{(i)}$ denote the i th order statistic of X_1, \dots, X_n , we would like to show that

$$\Pr [X_{(n)} - X_{(n-1)} > c] \geq \exp(\beta c).$$

Since X_i s are independent, the memoryless property of exponential distributions gives that when conditioned on

$X_i \geq 0$, X_i still follows an exponential distribution with mean $1/\beta$. For all subsets $S \subseteq \{1 \dots n\}$, let \mathcal{E}_S denote the event that for all $i \in S$, $X_i \geq 0$, and for all $i \notin S$, $X_i < 0$. By the law of total probability, we have

$$\begin{aligned} & \Pr [X_{(n)} - X_{(n-1)} > c] \\ &= \sum_S \Pr [X_{(n)} - X_{(n-1)} > c \mid \mathcal{E}_S] \Pr [\mathcal{E}_S]. \end{aligned}$$

Since $X_n = \delta_n \geq 0$, $\Pr [\mathcal{E}_S] = 0$ when $S = \emptyset$ or $S \not\ni n$. The only other case with $|S| = 1$ is when $S = \{n\}$. Here we have $\Pr [X_1 > c] \geq 1 - \exp(-\beta c)$. Combining this with $X_{(n)} \geq X_1$ and $X_{(n-1)} < 0$ gives a probability of at least $\exp(\beta c)$.

It remains to consider the case where $|S| \geq 2$. In this case both $X_{(n)}$ and $X_{(n-1)}$ are from elements in S , so it suffices to consider the X_i s given by $i \in S$. These $|S|$ variables are distributed the same as $|S|$ independent random variables following $\text{Exp}(\beta)$. Therefore by the distribution of order statistics given in Fact 3.1 we have:

$$\Pr [X_{(n)} - X_{(n-1)} > c \mid \mathcal{E}_S, |S| \geq 2] = \exp(-\beta c).$$

This means for any S , we have

$$\Pr [X_{(n)} - X_{(n-1)} > c \mid \mathcal{E}_S] \Pr [\mathcal{E}_S] \geq \exp(-\beta c) \Pr [\mathcal{E}_S].$$

Summing over all S and using the fact that $\sum_S \Pr [\mathcal{E}_S] = 1$ gives that

$$\Pr [X_{(n)} - X_{(n-1)} > c] \geq \exp(-\beta c),$$

or equivalently

$$\Pr [X_{(n)} - X_{(n-1)} \leq c] \leq 1 - \exp(-\beta c) < \beta c. \quad \blacksquare$$

Using this Lemma with $c = 1$ and applying linearity of expectation gives the bound on the number of edges between pieces.

COROLLARY 4.5. *The probability of an edge $e = uv$ having u and v in different pieces is bounded by $O(\beta)$, and the expected number of edges between pieces is $O(\beta m)$.*

5. IMPLEMENTATION AND PARALLELIZATION

Our partition routine as described in Algorithm 2 requires computing $\text{dist}_{-\delta}(u, v)$ for all pairs of vertices u and v . Standard modifications allow us to simplify it to the form shown in Algorithm 1, which computes BFS involving small integer distances.

The first observation is that the $-\delta_u$ shift at vertex u can be simulated by introducing a super source s with distance $-\delta_u$ to each vertex u . Then if we compute single source shortest path from s to all vertices, the component that v belongs to is given by the first vertex on the shortest path from s to it. Two more observations are needed to transform this shortest path setup to a BFS. First, the negative lengths on edges leaving s can be fixed by adding $\delta_{\max} = \max_u \delta_u$ to all these weights. Second, note that the only edges with non-integer lengths are the ones leaving s . In this shortest path algorithm, the only time that we need to examine the non-integer parts of lengths is when we compare two distances whose integer parts are tied. So the fractional parts can be viewed as tie-breakers for equal integer distances, and all

distances with the same integer part can be processed in parallel. We'll show below that these tie breakers can also be replaced by a random permutation of integers.

Therefore, the algorithm is equivalent to computing shortest path when all edge lengths are integer, with an extra tie breaking rule for comparing distances. In order to use unweighted BFS, it remains to handle the edges with non-unit lengths leaving s , and we do so by processing the those edges in a delayed manner. An edge from s to v only causes v to be added to the BFS queue when the frontier of the search has reached a distance larger than the length of that edge and v has not been visited yet. So it suffices to check all vertices v with a length L edge to s when the BFS frontier moves to distance L , and add the unvisited ones to that level.

The exact cost of running a parallel breadth first search depends on the model of parallelism. There has been much practical work on such routines when the graph has small diameter [21, 8, 26]. For simplicity we will use the $O(\Delta \log n)$ depth and $O(m)$ work bound in the PRAM model given in [18]. Here Δ is the maximum distance that we run the BFS to, and can be bounded by $O(\frac{\log n}{\beta})$. This allows us to prove our main claim about the performance of PARTITION.

Proof of Theorem 1.2: Consider running PARTITION using the BFS based implementation described above, and repeating until we have an $(\beta, O(\frac{\log n}{\beta}))$ partition. Since the $\delta_{u,s}$ are generated independently, they can be computed in $O(n)$ work and $O(1)$ depth in parallel. The rest of the running time comes from assigning vertices to pieces using shifted shortest path. As the maximum distance from a vertex to the center of its piece is $O(\frac{\log n}{\beta})$ (or we could stop the algorithm at this point), this BFS can be done in $O(m)$ work and $O(\frac{\log^2 n}{\beta})$ depth using parallel BFS algorithms. The resulting decomposition can also be verified in $O(\log n)$ depth and $O(m)$ time.

It remains to bound the success probability of each iteration. Lemma 4.1 gives that the shortest path from u to any $v \in S_u$ is contained in S_u , so $\max_{v \in S_u} \text{dist}(u, v)$ is an upper bound for the strong diameter of each subset. For each vertex $v \in S_u$, since $\text{dist}_{-\delta}(v, v) = d(v, v) - \delta_v \leq 0$ is a candidate, $\text{dist}_{-\delta}(u, v) \leq -\delta_v$. Lemma 4.2 then allows us to bound this value by $O(\log n/\beta)$ with high probability. The expected number of edges between pieces follows from Corollary 4.5, so with constant probability we meet both requirements of a $(\beta, O(\frac{\log n}{\beta}))$ partition. Therefore, we are expected to iterate a constant number of times, giving the expected depth and work bounds. \blacksquare

One practical aspect worth noting is that the fractional parts of the δ values can be viewed as a lexicographical ordering upon all vertices which are used for tie breaking. This is where the tie breaking rule specified in Section 4 may be of use. As the exponential distribution is memoryless and the shifts are generated independently, the fractional parts can also be emulated by directly generating a random permutation of the vertices. This view is perhaps closer to the use of random permutations in the optimal tree-metric embedding algorithm [16].

Similar ideas may also be used in practice instead of computing δ_u . Although generating random variables from such distributions have been studied extensively [19], avoiding these routines might further reduce the cost of this stage of the algorithm. One possibility is to generate a random permutation of the vertices, and assign the shift values based

on positions in the permutation. We believe that the slight changes in distributions could be accounted for using a more intricate analysis, but might be more easily studied empirically.

6. CONCLUSION / REMARKS

We showed a simple parallel algorithm for computing low diameter decompositions of undirected unweighted graphs. Given a graph G with n vertices and m edges along with any parameter β , it returns a $(\beta, O(\frac{\log n}{\beta}))$ decomposition in $O(\frac{\log^2 n}{\beta})$ depth and $O(m)$ work. This routine can be used in place of PARTITION from [9] to give a faster algorithm for solving SDD linear systems. It also represents a different view of ball growing, which is at the core of the best sequential low stretch spanning tree algorithms [15, 1, 2].

We believe that our approach may lead to a variety of improvements in algorithms that use ball growing or low diameter decompositions. Many of these applications take place in the weighted setting, and rely on additional clustering-based properties. As a result, many of them are perhaps better examined on a per-application basis. The analysis of the partition routine from Section 4 can be readily extended to the weighted case. However, the depth of the algorithm is harder to control since hop count is no longer closely related to diameter. We believe obtaining similar parallel guarantees in the weighted setting, as well as showing clustering-based properties are interesting directions for future work.

7. REFERENCES

- [1] I. Abraham, Y. Bartal, and O. Neiman. Nearly tight low stretch spanning trees. pages 781–790, 2008.
- [2] I. Abraham and O. Neiman. Using petal-decompositions to build a low stretch spanning tree. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 395–406, New York, NY, USA, 2012. ACM.
- [3] N. Alon, R. Karp, D. Peleg, and D. West. A graph-theoretic game and its application to the k -server problem. *SIAM J. Comput.*, 24(1):78–100, 1995.
- [4] B. Awerbuch. Complexity of network synchronization. *J. Assoc. Comput. Mach.*, 32(4):804–823, 1985.
- [5] B. Awerbuch, B. Berger, L. Cowen, and D. Peleg. Low-diameter graph decomposition is in NC. In *Proceedings of the Third Scandinavian Workshop on Algorithm Theory*, SWAT '92, pages 83–93, London, UK, UK, 1992. Springer-Verlag.
- [6] B. Awerbuch, A. V. Goldberg, M. Luby, and S. A. Plotkin. Network decomposition and locality in distributed computation (extended abstract). pages 364–369, 1989.
- [7] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on*, pages 184–193, 1996.
- [8] S. Beamer, K. Asanović, and D. Patterson. Direction-optimizing breadth-first search. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, SC '12, pages 12:1–12:10, Los Alamitos, CA, USA, 2012. IEEE Computer Society Press.
- [9] G. Blelloch, A. Gupta, I. Koutis, G. Miller, R. Peng, and K. Tangwongsan. Nearly-linear work parallel sdd solvers, low-diameter decomposition, and low-stretch subgraphs. *Theory of Computing Systems*, pages 1–34, March 2013.
- [10] G. E. Blelloch, A. Gupta, and K. Tangwongsan. Parallel probabilistic tree embeddings, k -median, and buy-at-bulk network design. In *Proceedings of the 24th ACM symposium on Parallelism in algorithms and architectures*, SPAA '12, pages 205–213, New York, NY, USA, 2012. ACM.
- [11] P. Christiano, J. A. Kelner, A. Mądry, D. Spielman, and S.-H. Teng. Electrical Flows, Laplacian Systems, and Faster Approximation of Maximum Flow in Undirected Graphs. In *Proceedings of the 43rd ACM Symposium on Theory of Computing (STOC)*, 2011.
- [12] E. Cohen. Fast algorithms for constructing t -spanners and paths with stretch t . *SIAM J. Comput.*, 28(1):210–236, 1998.
- [13] E. Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest paths. *J. ACM*, 47(1):132–166, 2000.
- [14] S. I. Daitch and D. A. Spielman. Faster approximate lossy generalized flow via interior point algorithms. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, STOC '08, pages 451–460, New York, NY, USA, 2008. ACM.
- [15] M. Elkin, Y. Emek, D. A. Spielman, and S.-H. Teng. Lower-stretch spanning trees. *SIAM Journal on Computing*, 38(2):608–628, 2008.
- [16] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. System Sci.*, 69(3):485–497, 2004.
- [17] W. Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons Inc., New York, 1971.
- [18] P. N. Klein and S. Subramanian. A randomized parallel algorithm for single-source shortest paths. *Journal of Algorithms*, 25:205–220, 1997.
- [19] D. E. Knuth. *The art of computer programming, volume 2 (3rd ed.): seminumerical algorithms*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.
- [20] T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, 1999.
- [21] C. E. Leiserson and T. B. Scharidl. A work-efficient parallel breadth-first search algorithm (or how to cope with the nondeterminism of reducers). In *Proceedings of the 22nd ACM symposium on Parallelism in algorithms and architectures*, SPAA '10, pages 303–314, New York, NY, USA, 2010. ACM.
- [22] N. Linial and M. Saks. Decomposing graphs into regions of small diameter. In *Proceedings of the second annual ACM-SIAM symposium on Discrete algorithms*, SODA '91, pages 320–330, Philadelphia, PA, USA, 1991. Society for Industrial and Applied Mathematics.
- [23] S. Plotkin, S. Rao, and W. D. Smith. Shallow excluded minors and improved graph decompositions. In *Proceedings of the fifth annual ACM-SIAM symposium on Discrete algorithms*, SODA '94, pages

462–470, Philadelphia, PA, USA, 1994. Society for Industrial and Applied Mathematics.

- [24] J. Sherman. Breaking the multicommodity flow barrier for $O(\sqrt{\log n})$ -approximations to sparsest cut. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science, FOCS '09*, pages 363–372, Washington, DC, USA, 2009. IEEE Computer Society.
- [25] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:888–905, 1997.
- [26] J. Shun and G. E. Blelloch. Ligra: A lightweight graph processing framework for shared memory. 2013.
- [27] K. S. Trivedi. *Probability and statistics with reliability, queuing and computer science applications*. John Wiley and Sons Ltd., Chichester, UK, 2nd edition, 2002.
- [28] C. Wulff-Nilsen. Separator theorems for minor-free and shallow minor-free graphs with applications. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS '11*, pages 37–46, Washington, DC, USA, 2011. IEEE Computer Society.

APPENDIX

A. ALTERNATE PROOF OF KEY PARTITION LEMMA

We give alternate, more formulaic proofs of Lemma 4.2, which bounds the difference between the shortest and second shortest shifted distance to any point in the graph.

Proof of Lemma 4.4:

Let \mathcal{E} denote the number of indices i such that:

$$d_i - \delta_i \leq d_j - \delta_j + c \quad \forall j.$$

For each vertex i , let \mathcal{E}_i be an indicator variable for the event that:

$$d_i - \delta_i \leq d_j - \delta_j + c \quad \forall j.$$

We will integrate over the value of $t = d_i - \delta_i$. For a fixed value of t , \mathcal{E}_i occurs if and only if $\delta_j \leq d_j - t + c$ for each j . As the shift values are picked independently, we can multiply the cumulative distribution functions for $Exp(\beta)$ and get:

$$\begin{aligned} \Pr[\mathcal{E}_i] &= \int_{t=-\infty}^{\infty} f_{Exp}(d_i - t, \beta) \prod_{j \neq i} F_{Exp}(d_j - t + c, \beta). \end{aligned}$$

When $t > d_1 + c$, $d_1 - t + c < 0$ and $f_{Exp}(d_1 - t, \beta) = F_{Exp}(d_1 - t + c, \beta) = 0$. So it suffices to evaluate this integral up to $t = d_1 + c$. Also, we may use $\exp(-\beta x)$ as an upper bound as $f_{Exp}(x, \beta)$, and arrive at:

$$\begin{aligned} \Pr[\mathcal{E}_i] &\leq \int_{t=-\infty}^{d_1+c} \beta \exp(-\beta(d_i - t)) \prod_{j \neq i} F_{Exp}(d_j - t + c) \\ &\leq \int_{t=-\infty}^{d_1+c} \beta \exp(-\beta(d_i - t)) \prod_{j \neq i} (1 - \exp(-\beta(d_j - t + c))). \end{aligned}$$

We now bound $\mathbf{E}[\mathcal{E}] = \mathbf{E}[\sum_i \mathcal{E}_i]$. By linearity of expectation we have:

$$\begin{aligned} \mathbf{E}\left[\sum_i \mathcal{E}_i\right] &\leq \sum_i \int_{t=-\infty}^{d_1+c} \beta \exp(-\beta(d_i - t)) \\ &\quad \prod_{j \neq i} (1 - \exp(-\beta(d_j - (t - c)))) \\ &= \exp(\beta c) \int_{t=-\infty}^{d_1+c} \beta \sum_i \exp(-\beta(d_i - t + c)) \\ &\quad \prod_{j \neq i} (1 - \exp(-\beta(d_j - t + c))). \end{aligned}$$

Observe that the expression being integrated is the derivative w.r.t. t of:

$$-\prod_i (1 - \exp(-\beta(d_i - t + c))).$$

Therefore we get:

$$\mathbf{E}[\mathcal{E}] \leq -\exp(\beta c) \prod_i (1 - \exp(-\beta(d_i - t + c))) \Bigg|_{t=-\infty}^{t=d_1+c}$$

When $t \rightarrow -\infty$, $-\beta(d_i - t + c) \rightarrow -\infty$. Therefore $\exp(-\beta(d_i - t + c)) \rightarrow 0$, and the overall product tends to $-\exp(\beta c)$.

When $t = d_1 + c$, we have:

$$\begin{aligned} &-\exp(\beta c) \prod_i (1 - \exp(-\beta(d_i - (d_1 + c) + c))) \\ &= -\exp(\beta c) \prod_i (1 - \exp(-\beta(d_i - d_1))) \\ &\leq -\exp(\beta c) \prod_i (1 - \exp(0)) = 0 \quad (\text{Since } d_i \geq d_1) \end{aligned}$$

Combining these two gives $\mathbf{E}[\mathcal{E}] \leq \exp(\beta c)$.

By Markov's inequality the probability of there being another vertex being within c of the minimum is at most $\exp(\beta c) - 1 \leq O(\beta c)$ for $c = 1$. \blacksquare