# Approximating the Statistics of various Properties in Randomly Weighted Graphs

Yuval Emek * Amos Korman † Yuval Shavitt ‡

**Abstract**

Consider the setting of *randomly weighted graphs*, namely, graphs whose edge weights are chosen independently according to probability distributions with finite support over the non-negative reals. Under this setting, weighted graph properties such as the *diameter*, the *radius* (with respect to a designated vertex), and the weight of a *minimum spanning tree* become random variables and we are interested in computing their expectation. Unfortunately, this turns out to be #P-hard. In this paper, we define a family of weighted graph properties (that includes the above three) and show that for each property in this family, the problem of computing the $k^{th}$ *moment* (and in particular, the expectation) of the corresponding random variable admits a *fully polynomial-time randomized approximation scheme (FPRAS)* for every fixed $k$.

## 1 Introduction

In the traditional setting of graph algorithms, the input is typically a graph $G$ whose edges are often associated with some *weights*. In most applications, $G$ represents a real-life network and the edge weights correspond to some attributes of the network's links which are assumed to be known when one wishes to apply some graph algorithm to $G$. Unfortunately, in many scenarios these attributes of the real-life network cannot be determined. Still, however, it is often believed that the attributes of the network's links are governed by some known probability distributions. For example, the latency along each communication link in the Internet is usually assumed to be a random variable, rather than a fixed value that can be determined a priori; Internet tomography techniques are devised to reveal the distributions of these random variables [13, 11]. The random delay behavior also applies to overlay networks (where each link corresponds to a path in the Internet) and a recent Internet measurement indicates that some path delays exhibit strict bi-modal characteristics [36]. A similar situation occurs in the emerging field of networking called delay tolerant networks (DTNs) [15] that includes sparse ad-hoc networks [41, 4], space exploration networks [9], submarine networks [35], and sensor networks.

Our goal in this paper is to advance the theoretic foundations of graph algorithms operating in the context of edge weights that obey some specified probability distributions. In this context, various properties of the graph become random variables and we wish to design better algorithms for the problems of computing the statistical features of these random variables. This turns out to be a non-trivial task even for basic and fundamental graph properties which are easy to compute in the traditional (deterministic) setting. Similar challenges were previously tackled in many works (see the survey of Ball et al. [3]), however the computational angle of these problems received little treatment.

**The model.** A *randomly weighted (RW) graph* is a graph[1] $G$ in which the edge weights are independent random variables with finite support over the non-negative reals. Specifically, every edge $e \in E(G)$ is associated with some positive integer $m(e)$ and with some non-negative reals $W_e^1, \ldots, W_e^{m(e)}$ and $p_e^1, \ldots, p_e^{m(e)}$, where $\sum_{i=1}^{m(e)} p_e^i = 1$, such that the weight of $e$ is $w(e) = W_e^i$ with probability $p_e^i$ independently of all other edges. The reals $W_e^1, \ldots, W_e^{m(e)}$ are called the *phases* of edge $e$. An RW graph in which all the edge weights are identically distributed is referred to as an *identically distributed weighted graph*.

**Weighted graph properties.** Let $\mathcal{G}$ be the collection of all (not necessarily simple) connected graphs $G$ with (deterministic) non-negative edge weights $w : E(G) \to \mathbb{R}_{\geq 0}$. Throughout, we think of a *weighted graph property* as a function $\mathcal{X} : \mathcal{G} \to \mathbb{R}_{\geq 0}$ that assigns a non-negative real $\mathcal{X}(G)$ to each graph $G \in \mathcal{G}$. A weighted graph property $\mathcal{X}$ is said to be *distance-cumulative* if it satisfies the following requirements for every graph

*Microsoft Israel R&D Center, Herzliya, Israel.

†CNRS and Université Paris Diderot - Paris 7, France. Supported in part by the ANR project ALADDIN, by the INRIA project GANG, and by COST Action 295 DYNAMO.

‡School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel.

[1] Unless stated otherwise, all graphs in this paper are assumed to be finite and undirected, though not necessarily simple. The vertex set and edge set of graph $G$ are denoted by $V(G)$ and $E(G)$, respectively.

$G \in \mathcal{G}$:
**(R1)** If $G' \in \mathcal{G}$ is the graph obtained from $G$ by multiplying all edge weights by some factor $r \in \mathbb{R}_{\geq 0}$, then $\mathcal{X}(G') = r \cdot \mathcal{X}(G)$.
**(R2)** If $G' \in \mathcal{G}$ is the graph obtained from $G$ by increasing the weight of some edge $e \in E(G)$ by an additive term $r \in \mathbb{R}_{\geq 0}$, then $\mathcal{X}(G) \leq \mathcal{X}(G') \leq \mathcal{X}(G) + r$.
**(R3)** If $w(e) = 0$ for some edge $e \in E(G)$, then $\mathcal{X}(G/e) = \mathcal{X}(G)$, where $G/e \in \mathcal{G}$ is the graph obtained from $G$ by contracting the edge $e$ and removing self loops.
**(R4)** If $e, e' \in E(G)$ are parallel edges and $w(e) \geq w(e')$, then $\mathcal{X}(G - e) = \mathcal{X}(G)$, where $G - e \in \mathcal{G}$ is the graph obtained from $G$ by removing the edge $e$.
**(R5)** If $\mathcal{X}(G)$ is strictly positive, then $\mathcal{X}(G) \geq \min\{w(e) > 0 \mid e \in E(G)\}$.
**(R6)** $\mathcal{X}(G) = 0$ if and only if the graph obtained from $G$ by removing all (strictly) positive weight edges is connected.

Requirements (R1)–(R5) and the sufficiency direction of requirement (R6) are naturally satisfied by weighted graph properties in which the edge weights correspond to time delays, routing costs, etc. The necessity direction of requirement (R6) is slightly more restrictive. We extend the definition of distance-cumulative weighted graph properties $\mathcal{X}$ by assuming that $\mathcal{X}(G) = \infty$ for every disconnected graph $G$. A weighted graph property $\mathcal{X}$ is said to be *efficiently calculated* if there exists a polynomial time algorithm that computes $\mathcal{X}(G)$ for every (deterministically weighted) graph $G$.

Consider some graph $G \in \mathcal{G}$. The *diameter* of $G$, denoted $\mathrm{diam}(G)$, is defined as $\mathrm{diam}(G) = \max\{\mathrm{dist}_G(u, v) \mid u, v \in V(G)\}$, where $\mathrm{dist}_G(u, v)$ is the *distance* between $u$ and $v$ in $G$, namely, the length of a shortest (with respect to the edge weights) path from $u$ to $v$. For a designated vertex $u \in V(G)$, the *radius* of $G$, denoted $\mathrm{rad}(G)$, is defined as $\mathrm{rad}(G) = \max\{\mathrm{dist}_G(u, v) \mid v \in V(G)\}$. The weight of a *minimum spanning tree* of $G$, denoted $\mathrm{MST}(G)$, is defined as $\mathrm{MST}(G) = \min\{\sum_{e \in T} w(e) \mid T \text{ is a spanning tree of } G\}$. It is easy to verify that $\mathrm{diam}(G)$, $\mathrm{rad}(G)$, and $\mathrm{MST}(G)$ are distance-cumulative weighted graph properties. Efficient algorithms that compute them are described in most textbooks on graph algorithms (e.g., [14]). Another distance-cumulative weighted graph property is the diameter of a *min-diameter spanning tree*, defined as $\min\{\mathrm{diam}(T) \mid T \text{ is a spanning tree of } G\}$. The min-diameter spanning tree is also known to be efficiently calculated [21].

Two more efficiently calculated weighted graph properties that fit into the definition of distance-cumulative properties by slightly modifying requirement (R5) are the *all pairs average distance* and the *single source average distance* with respect to some designated vertex. Our techniques can be adjusted to handle such a modification, however, this is beyond the scope of the current version of the paper.

**Network reliability.** A setting closely related to RW graphs is that of *network reliability*. A network reliability instance is depicted by a graph $G$ (hereafter, a *network reliability graph*) in which every edge $e \in E(G)$ is associated with a *failure* probability $0 \leq p_e < 1$. It is assumed that each edge $e$ *fails*, i.e., removed from the graph, with probability $p_e$ independently of all other edges. With regards to the network reliability setting, one is usually interested in the probability that the graph remains connected, referred to as the *all-terminal network reliability (ATNR)*), or in the probability that two designated vertices remain in the same connected component, referred to as the *two-terminal network reliability (TTNR)*.

We shall make an extensive use of the notion of a *weighted network reliability (WNR) graph* which is the weighted counterpart of network reliability graphs: In a WNR graph $G$ every edge $e \in E(G)$ is associated with a non-negative real weight $w(e)$ apart from the failure probability $0 \leq p_e < 1$. In this context we will usually be interested in the random variable $\mathcal{X}(G)$, where $\mathcal{X}$ is some weighted graph property.

**Our contribution.** Let $\mathcal{X}$ be some distance-cumulative weighted graph property and assume that it is efficiently calculated. Given some connected RW graph $G$, $\mathcal{X}(G)$ is a random variable — denote it by $X$ — and we are interested in approximating its $k^{th}$ *moment*, i.e., $\mathbb{E}[X^k]$, for any fixed $k \geq 1$. Specifically, we develop a *fully polynomial time randomized approximation scheme (FPRAS)* for the problem, namely, a randomized algorithm that runs in time $\mathrm{poly}(|G|, 1/\epsilon)$ for any choice of $\epsilon > 0$, where $|G|$ stands for the number of bits required to encode $G$ in a standard binary representation[2], and returns a $(1+\epsilon)$-approximation of $\mathbb{E}[X^k]$ with probability[3] at least 3/4. The following theorem is established in Sections 3, 4, and 5.

THEOREM 1.1. *The problem of computing the $k^{th}$ moment of $\mathcal{X}(G)$ on connected RW graphs $G$ admits an FPRAS for every fixed $k$.*

In general, Theorem 1.1 is best possible. This is because exact solutions to the problems of computing $\mathbb{E}[\mathrm{diam}(G)]$ and $\mathbb{E}[\mathrm{rad}(G)]$ are #P-hard to obtain even

[2]The running time of our FPRAS is measured with respect to $|G|$, rather than $|V(G)| + |E(G)|$, hence it is not strongly polynomial.

[3]Using the *median of means method*, the success probability of an FPRAS can be increased to $1 - \hat{\epsilon}$ for any choice of $\hat{\epsilon} > 0$ at the cost of increasing the run-time by an $O(\log(1/\hat{\epsilon}))$ factor.

when the input is restricted to identically distributed weighted graphs. Refer to Section 6 for a proof of this rather simple observation.

To the best of our knowledge, our FPRAS yields the first provably polynomial time algorithm with guaranteed approximation ratio for any non-trivial statistical feature of a weighted graph property in randomly weighted graphs. Moreover, it seems that most related literature focuses on individual weighted graph properties and does not attempt to provide a framework for a more general theory of such properties; indeed, Snyder & Steele call for such a framework in their survey [37]. We hope that our technique which is suitable for all distance-cumulative weighted graph properties will be a significant step in that direction.

**Related work.** The algorithmic aspects of randomly weighted graphs have been extensively studied since the early 60's (cf. Fulkerson [19]) mainly in the context of the shortest $(s,t)$-path, the longest $(s,t)$-path (a.k.a. the PERT problem), and maximum $(s,t)$-flow. A comprehensive account of the various methods developed for the computation (and approximation) of the statistical features corresponding to these weighted graph properties is provided by Ball et al. [3] who also observe that an exact computation of the expected values of these weighted graph properties[4] is #P-hard. Note that except for a few special cases (e.g., series-parallel networks with a specific type of edge distributions), none of the algorithms developed in that context is provably polynomial with guaranteed approximation ratio.

Distance-cumulative weighted graph properties were also investigated under this setting. Hassin & Zemel [22] prove that the diameter (and radius) of a complete $n$-vertex graph whose edge weights are uniformly and independently distributed in $[0,1]$ is almost surely $\Theta(\log(n)/n)$. The hidden constants in this expression were resolved by Janson [25] who shows that the $(s,t)$-distance, radius with respect to $s$, and diameter of a complete $n$-vertex graph whose edge weights are uniformly and independently distributed in $[0,1]$ converge in probability to $\ln n/n$, $2\ln n/n$, and $3\ln n/n$, respectively. Frieze [16] shows that for every distribution function $F$ with finite variance whose derivative at zero exists and satisfies $F'(0) = D > 0$, if the edge weights in a complete $n$-vertex graph $G$ are independently distributed according to $F$, then the weight of a minimum spanning tree of $G$ converges in probability to $\zeta(3)/D$, where $\zeta(3) = \sum_{j=1}^{\infty} 1/j^3$. This is generalized by Steele [38] who shows that the assumption on the finite variance can be lifted. For the special case of $F$ being the uniform distribution over $[0,1]$, Beveridge et al. [7] establish a variant of this bound for $r$-regular graphs.

Asymptotic results for minimum spanning trees on $n$ points uniformly and independently distributed in the Euclidean unit ball are established by Bertsimas & van Ryzin [6]. Kulkarni [31] and Alexopoulos & Jacobson [1] present algorithms that compute the distribution of $\mathrm{MST}(G)$ for graphs $G$ whose edge weights obey exponential and discrete distributions, respectively. The run-times of their algorithms are not necessarily polynomial, though. A non-trivial upper bound on $\mathbb{E}[\mathrm{MST}(G)]$ is established by Jain & Mamer [24].

The network reliability setting was introduced by Moore and Shannon [34]; refer to [3] for a comprehensive survey. Valiant [39] showed that the exact computation of both two-terminal and all-terminal network reliability are #P-hard. An FPRAS for the probability that a network reliability graph becomes disconnected following an edge failure event (that is, for the probability that complements ATNR) is developed by Karger [26]. As described later on, some of the ingredients of Karger's technique are fundamental to the development of our FPRAS.

A different, yet, related subject that admits a plethora of literature is average case analysis for graph algorithms (e.g. [17, 27, 12, 32]). There it is assumed that the edge weights in the input of some graph algorithm are drawn from a specified probability distribution and the goal is to analyze the expected run-time of the algorithm with respect to that distribution; refer to [18] for a survey. It is important to point out that upon invocation of the graph algorithm, the actual edge weights are determined, and in particular, known to the algorithm, in contrast to the RW graphs setting in which the challenge is to cope with the uncertainty in the edge weights.

**Techniques.** We now provide an informal overview of the construction of an FPRAS for $\mathbb{E}[\mathcal{X}(G)]$, where $\mathcal{X}$ is some efficiently calculated distance-cumulative weighted graph property and $G$ is an RW graph; approximating higher moments is very similar. Note that if the variance of $\mathcal{X}(G)$ is at most some polynomial times $\mathbb{E}[\mathcal{X}(G)]^2$ (that is, the critical ratio is polynomial), then $\mathbb{E}[\mathcal{X}(G)]$ can be approximated by means of direct sampling (a.k.a. Monte Carlo method). However, there exist some simple examples (cf. Appendix A) in which the variance is too large, and therefore a different approach must be sought.

The desired approximation would have been straightforward to obtain if we could have efficiently approximated $\mathbb{P}(\mathcal{X}(G) > x)$ for arbitrary choices of real

[4] In Section 6 we establish a similar hardness result for the diameter and radius which are not among the weighted graph properties covered by Ball et. al. [3].

$x \geq 0$. Unfortunately, obtaining such an approximation to within a multiplicative error of $1 + O(\epsilon)$ seems to be a challenging task. In fact, when $\mathcal{X} = \text{MST}$ (namely, we are required to approximate the expected weight of a minimum spanning tree), an efficient implementation of this task would yield an FPRAS for the Tutte polynomial $T_G(x, y)$ of arbitrary graphs $G$ for every $x, y > 1$ (refer to Bollobás [8] for a comprehensive treatment of the Tutte polynomial and its many applications); whether or not such an FPRAS exists is an important open question [2, 20].

Instead, we use a careful "sliding window" argument to show that the desired approximation of $\mathbb{E}[\mathcal{X}(G)]$ can be efficiently obtained by repeatedly invoking a procedure called Procedure `Estimate`: given a positive real $x$, Procedure `Estimate` approximates $\mathbb{P}(\mathcal{X}(G) > x)$ to within an additive error of $O(\epsilon) \cdot \mathbb{P}(\mathcal{X}(G) > 0)$. In other words, we show that it suffices to produce a weaker approximation of $\mathbb{P}(\mathcal{X}(G) > x)$; the quality of this weaker approximation is determined by $\mathbb{P}(\mathcal{X}(G) > 0)$. The "sliding window" argument, presented in Section 3, is based on iteratively setting to 0 all edge phases in $G$ smaller than some threshold for a carefully chosen sequence of thresholds.

Procedure `Estimate` itself is presented in Section 4. As a preliminary step, we transform the RW graph $G$ into a WNR graph $H$ (see Section 5), so that the challenge becomes to approximate $\mathbb{P}(\mathcal{X}(H) > x)$ to within an additive error of $O(\epsilon) \cdot \mathbb{P}(\mathcal{X}(H) > 0)$ for a given WNR graph $H$.[5] To that end, we consider the subgraph $H^0$ consisting of the zero weight edges of $H$ and formulate the real valued random variable $Y$ that maps each instance $\mathcal{I}$ of the probability space defined by $E(H) - E(H^0)$ to $\mathbb{P}(\mathcal{X}(H) > x \mid \mathcal{I})$. Since $\mathbb{E}[Y] = \mathbb{P}(\mathcal{X}(H) > x)$, it is sufficient to approximate $\mathbb{E}[Y]$. We would have wanted to do so by sampling instances $\mathcal{I}$ of the probability space defined by $E(H) - E(H^0)$ and then computing $\mathbb{P}(\mathcal{X}(H) > x \mid \mathcal{I})$. Sampling instances of the probability space defined by $E(H) - E(H^0)$ is a straightforward task. The problem is that given such an instance $\mathcal{I}$, it is not clear how to efficiently compute $\mathbb{P}(\mathcal{X}(H) > x \mid \mathcal{I})$.

To tackle this obstacle, we revisit Karger's FPRAS for the ATNR problem [26]. Karger's technique is based on identifying a collection $\mathcal{C}$ of polynomially many (2-way) cuts in $H^0$ such that the probability that all edges of some cut not in $\mathcal{C}$ fail is small. For each instance $\mathcal{I}$, we identify those cuts in $\mathcal{C}$ that conditioned on $\mathcal{I}$, imply $\mathcal{X}(H) > x$. For that to work, we must extend Karger's construction of $\mathcal{C}$ to $r$-way cuts for all $r \geq 2$. This extension builds upon the recent Bell number bound of Berend & Tassa [5]. We then employ the method of Karp, Luby, and Madras [29, 30] for probabilistic DNF satisfiability to approximate the probability that at least one of these cuts is induced by the failing edges in $E(H^0)$.

[5] Approximating $\mathbb{P}(\mathcal{X}(H) > x)$ to within an additive error of $O(\epsilon) \cdot \mathbb{P}(\mathcal{X}(H) > 0)$ can be done via another method. This method, which was suggested to us by an anonymous reviewer, is simpler than the one we describe in Section 4, however, it relies on randomly sampling satisfying assignments to a DNF formula. Although such sampling can be done by employing the method of Karp, Luby, and Madras [29, 30], the time bound guarantee of the resulting randomized algorithm is polynomial only on expectation, and thus does not yield a "proper" FPRAS. This is discussed in more detail in the full version of the paper.

## 2 Preliminaries

**Randomly weighted graphs.** Throughout we consider some distance-cumulative weighted graph property $\mathcal{X}$. Let $G$ be some $n$-vertex connected RW graph and let $X$ denote the random variable that takes on $\mathcal{X}(G)$. By requirement (R1), we may assume without loss of generality that the edge phases of $G$ are scaled so that the smallest non-zero phase is exactly 1. Consequently requirement (R5) implies that $X$ is either 0, or it is bounded from below by 1. On the other hand, requirements (R2) and (R4) guarantee that $X$ is bounded from above by $x_{\max} = n^2 \cdot \max\{W_e^i \mid e \in E(G) \text{ and } 1 \leq i \leq m(e)\}$.

**Weighted network reliability graphs.** Let $G$ be some WNR graph and let $F \subseteq E(G)$ be some edge subset. Recall that each edge $e \in F$ fails with probability $p_e$ — this defines a probability space. It will be convenient to view an *instance* $\mathcal{I}$ of this probability space as a Boolean function $\mathcal{I} : F \to \{0, 1\}$, where

$$\mathcal{I}(e) = \begin{cases} 0 & \text{if } e \text{ fails;} \\ 1 & \text{otherwise.} \end{cases}$$

At the risk of abusing notation, we may sometime write $F$ when we actually refer to the probability space it defines; our intentions will be clear from the context.

**Cuts and compact cuts.** Consider some connected graph $G$. An *$r$-way cut* $C$ of $G$ is a partition of $V(G)$ to $r$ pairwise disjoint subsets, that is, $C = \{U_1, \ldots, U_r\}$, where $\bigcup_{1 \leq i \leq r} U_i = V(G)$ and $U_i \cap U_j = \emptyset$ for every $i \neq j$. The subsets $U_1, \ldots, U_r$ are referred to as the *clusters* of $C$. A *cut* refers[6] to an $r$-way cut for any $r \geq 2$.

Consider some $r$-way cut $C = \{U_1, \ldots, U_r\}$ of $G$. We say that an edge $e \in E(G)$ *crosses* $C$ if $e \in U_i \times U_j$ for some $i \neq j$. The set of edges crossing $C$ is denoted by $E(C)$. The cardinality $|E(C)|$ is referred to as the *size*

[6] In some literature, a cut refers to a 2-way cut, while an $r$-way cut for $r > 2$ is called a *multiway* cut. We do not make this distinction.

of $C$; if the edges of $G$ are assigned with positive *costs* $c : E(G) \to \mathbb{R}_{>0}$, then the sum $\sum_{e \in E(C)} c(e)$ is referred to as the *cost* of $C$. The cut $C$ is called *compact* if $G(U_i)$ is connected for every $1 \leq i \leq r$. Note that every $r$-way cut is a compact $r'$-way cut for some $r' \geq r$. A *min cut* (respectively, *min cost cut*) is a cut of minimum size (resp., cost). It is easy to verify that a min cut (resp., min cost cut) must be a compact 2-way cut.

Consider some subset $F \subseteq E(G)$ and a compact cut $C$ of $G$. We say that $F$ *induces* the cut $C$ if the connected components of the graph obtained from $G$ by removing the edges in $F$ agree with the clusters of $C$. In particular, $F$ must be a superset of $E(C)$; it may contain additional edges as long as the removal of these edges does not disconnect any cluster of $C$.

**Monte Carlo method and approximators.** Consider some probability space with sample space $\Omega$ and let $X : \Omega \to \mathbb{R}$ be a real valued random variable over this probability space. Suppose that the expectation of $X$ is defined and denote it by $\mu$. Let $X_1, \ldots, X_n$ be $n$ independent samples of $X$ and fix $\bar{X} = \sum_{i=1}^{n} X_i/n$. Evaluating $\mu$ by $\bar{X}$ is referred to as the *Monte Carlo method* (cf. [29]). Let $\epsilon$ and $\hat{\epsilon}$ be some positive reals. The following two theorems are direct consequences of Chernoff's inequality [10] (Theorem 2.1) and Hoeffding's inequality [23] (Theorem 2.2).

THEOREM 2.1. *If $X$ is an indicator random variable (namely, $X \in \{0,1\}$), then taking $n \geq 4\ln(2/\hat{\epsilon})\mu/\epsilon^2$ samples guarantees that $\mathbb{P}(|\bar{X} - \mu| > \epsilon) \leq \hat{\epsilon}$.*

THEOREM 2.2. *If $X$ is almost surely in the interval $[a,b]$, where $b - a = \rho$, then taking $n \geq \ln(2/\hat{\epsilon})\rho^2/(2\epsilon^2)$ samples guarantees that $\mathbb{P}(|\bar{X} - \mu| > \epsilon) \leq \hat{\epsilon}$.*

This leads us to the notion of *approximators*. Consider some non-negative real value $v$ that we would like to approximate with the real $v'$. Then $v'$ is said to be an $(\epsilon, \hat{\epsilon})$-*approximator* of $v$ if it satisfies the inequality $|v - v'| \leq \epsilon$ with probability at least $1 - \hat{\epsilon}$, where the probability is taken over the randomness used to generate $v'$. Under this terminology, Theorems 2.1 and 2.2 provide sufficient conditions to guarantee that $\bar{X}$ is an $(\epsilon, \hat{\epsilon})$-approximator of $\mu$.

PROPOSITION 2.1. *If $v_1$ is an $(\epsilon_1, \hat{\epsilon_1})$-approximator of $v_0$ and $v_2$ is an $(\epsilon_2, \hat{\epsilon_2})$-approximator of $v_1$, then $v_2$ is an $(\epsilon_1 + \epsilon_2, \hat{\epsilon_1} + \hat{\epsilon_2})$-approximator of $v_0$.*

PROPOSITION 2.2. *If $v'_i$ is an $(\epsilon, \hat{\epsilon})$-approximator of $v_i$ for every $1 \leq i \leq n$, then $\sum_{i=1}^{n} v'_i/n$ is an $(\epsilon, n \cdot \hat{\epsilon})$-approximator of $\sum_{i=1}^{n} v_i/n$.*

## 3 An FPRAS for RW graphs

In this section we consider some $n$-vertex RW graph $G$ and some small performance parameter $\epsilon > 0$; our goal is to approximate $\mathbb{E}[X^k]$ to within a multiplicative error of $1 + O(\epsilon)$. Here, we restrict our attention to the case $k = 1$, that is, we approximate $\mathbb{E}[X]$. Extending our result to larger (yet fixed) values of $k$ is mainly a matter of notation and we omit it from this version of the paper. The approximation presented in this section builds upon the more sophisticated Procedure `Estimate` which is presented in Section 4.

THEOREM 3.1. *There exists a randomized algorithm that with probability at least $3/4$, approximates $\mathbb{E}[X]$ to within a multiplicative error of $1 + O(\epsilon)$ in time $\mathrm{poly}(|G|, 1/\epsilon)$.*

Let $N$ be the smallest integer such that $x_{\max} < (1+\epsilon)^N$. (Note that $N$ is proportional to $\log(x_{\max})/\epsilon = \mathrm{poly}(|G|, 1/\epsilon)$). Clearly, we have $0 \leq X < (1+\epsilon)^N$. Fix $\pi_i = \mathbb{P}(X \geq (1+\epsilon)^i)$ for every $0 \leq i \leq N$. Towards the approximation of $\mathbb{E}[X]$, we first define

$$\begin{aligned}
\mathcal{A} &= \sum_{i=1}^{N} (1+\epsilon)^{i-1} \cdot \mathbb{P}\left((1+\epsilon)^{i-1} \leq X < (1+\epsilon)^i\right) \\
&= \sum_{i=1}^{N} (1+\epsilon)^{i-1} \cdot [\pi_{i-1} - \pi_i] \\
&= \pi_0 + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \pi_i - (1+\epsilon)^{N-1} \cdot \pi_N \\
&= \pi_0 + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \pi_i \ , \qquad (3.1)
\end{aligned}$$

where (3.1) is due to the fact that $\pi_N = 0$. It is easy to verify that

$$\mathbb{E}[X]/(1+\epsilon) \leq \mathcal{A} \leq \mathbb{E}[X] \ , \qquad (3.2)$$

so our next goal is to approximate $\mathcal{A}$. Note that (3.1) enables the computation of a $(1+\epsilon)$-approximation of $\mathbb{E}[X]$ based on $(1+\epsilon)$-approximations of $\mathbb{P}(X \geq x)$ for sufficiently many values of $x$. Unfortunately, we do not know how to obtain such an approximation directly and we are forced to apply some modifications to $G$.

**The shrunk graphs.** Fix $\kappa = \left\lceil \log_{1+\epsilon}\left(\frac{\binom{n}{2}(1+\epsilon)}{\epsilon}\right)\right\rceil$. For $i = 0, 1, \ldots, N-1$, let $G_i$ be the RW graph obtained from $G$ by setting $W_e^{\varphi} \leftarrow 0$ for every edge $e \in E(G)$ and phase $1 \leq \varphi \leq m(e)$ such that $W_e^{\varphi} < (1+\epsilon)^{i-\kappa}$. We refer to a phase that was set to 0 in this process as a *shrunk* phase; the graphs $G_0, \ldots, G_{N-1}$ are called the *shrunk* graphs.

Let $X_i$ be the random variable that takes on $\mathcal{X}(G_i)$. Requirement (R2) guarantees that $X_0 \geq X_1 \geq \cdots \geq X_{N-1}$. The assumption that the minimum positive

phase is scaled to 1 implies that $G = G_0 = G_1 = \cdots = G_\kappa$, hence $X = X_0 = X_1 = \cdots = X_\kappa$. If $i > \kappa$, then $X_i$ may be smaller than $X$, however it is not much smaller as depicted in the following proposition.

PROPOSITION 3.1. *If $X \geq (1+\epsilon)^i$ for some $0 \leq i \leq N-1$, then $X/(1+\epsilon) < X_i \leq X$.*

*Proof.* Fix some instance $\mathcal{I}$ of the probability space defined by $E(G)$. By (R4), we may assume that the (deterministically) weighted graph resulting from $\mathcal{I}$ contains at most $\binom{n}{2}$ edges. Since phase $W_e^\varphi$ shrinks in $G_i$ only if $W_e^\varphi < (1+\epsilon)^{i-\kappa} \leq (1+\epsilon)^i \cdot \frac{\epsilon}{\binom{n}{2}(1+\epsilon)}$, (R2) guarantees that, subject to $\mathcal{I}$,

$$X_i > X - \binom{n}{2} \cdot \frac{\epsilon(1+\epsilon)^i}{\binom{n}{2}(1+\epsilon)} > X - \frac{\epsilon X}{1+\epsilon} = X/(1+\epsilon) \ .$$

The assertion follows as this argument holds for every instance $\mathcal{I}$.

Fix $\pi'_i = \mathbb{P}(X_i \geq (1+\epsilon)^i)$ for every $0 \leq i \leq N-1$ and define

$$\mathcal{A}' = \pi'_0 + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \pi'_i \ .$$

Clearly, $\pi'_i \leq \pi_i$ for every $0 \leq i \leq N-1$, thus $\mathcal{A}' \leq \mathcal{A}$. On the other hand, for every $1 \leq i \leq N-1$, we have

$$\begin{aligned} \pi'_{i-1} &= \mathbb{P}(X_{i-1} \geq (1+\epsilon)^{i-1}) \\ &\geq \mathbb{P}(X_i \geq (1+\epsilon)^{i-1}) \\ &\geq \mathbb{P}(X \geq (1+\epsilon)^i) \ = \ \pi_i \ , \end{aligned}$$

where the last inequality is due to Proposition 3.1. Since, $\pi'_0 = \pi_0$, we get

$$\begin{aligned} \frac{\mathcal{A}}{(1+\epsilon)} \ &= \ \frac{\pi_0}{(1+\epsilon)} + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-2} \cdot \pi_i \\ &\leq \ \frac{\pi'_0}{(1+\epsilon)} + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-2} \cdot \pi'_{i-1} \\ &= \ \pi'_0 + \sum_{i=2}^{N-1} \epsilon \cdot (1+\epsilon)^{i-2} \cdot \pi'_{i-1} \\ &= \ \pi'_0 + \sum_{i=1}^{N-2} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \pi'_i \ \leq \ \mathcal{A}' \ , \end{aligned}$$

therefore

$$\mathcal{A}/(1+\epsilon) \leq \mathcal{A}' \leq \mathcal{A} \ . \tag{3.3}$$

So, our next goal is to approximate $\mathcal{A}'$.

**Relying on local approximators.** Consider some $n$-vertex RW graph $H$ and denote the random variable that takes on $\mathcal{X}(H)$ by $X_H$. Let $\delta, \hat{\delta} > 0$ be some performance parameters. In Section 4 we present Procedure `Estimate` that given some $x > 0$, runs in time $\mathrm{poly}(n, 1/\delta, \log(1/\hat{\delta}))$ and returns a real $\mathcal{E}$ that serves as a $(\delta \cdot \mathbb{P}(X_H > 0), \hat{\delta})$-approximator of $\mathbb{P}(X_H \geq x)$. Set the performance parameters

$$\delta \leftarrow \frac{\epsilon}{(1+\epsilon)^\kappa} \approx (\epsilon/n)^2 \qquad \text{and} \qquad \hat{\delta} \leftarrow 1/(4N) \ .$$

For $i = 0, 1, \ldots, N-1$, we invoke Procedure `Estimate` on $H \leftarrow G_i$ with $x \leftarrow (1+\epsilon)^i$ to produce $\mathcal{E} \rightarrow \pi''_i$ which serves as a $(\delta \cdot \mathbb{P}(X_i > 0), \hat{\delta})$-approximator of $\pi'_i$.

Fix $\mathcal{A}'' = \pi''_0 + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \pi''_i$ and note that

$$\begin{aligned} |\mathcal{A}' - \mathcal{A}''| = \ & \left| \pi'_0 - \pi''_0 + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot (\pi'_i - \pi''_i) \right| \\ \leq \ & |\pi'_0 - \pi''_0| + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot |\pi'_i - \pi''_i| \ . \end{aligned}$$

By the choice of $\hat{\delta} = 1/(4N)$, we may use a union bound argument to conclude that the inequalities

$$|\pi'_i - \pi''_i| \leq \delta \cdot \mathbb{P}(X_i > 0)$$

hold simultaneously for all $0 \leq i \leq N-1$ with probability at least $3/4$; the remainder of this section is conditioned on that event. So,

$$\begin{aligned} & |\mathcal{A}' - \mathcal{A}''| \\ \leq \ & \delta \cdot \mathbb{P}(X_0 > 0) + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \delta \cdot \mathbb{P}(X_i > 0) \\ = \ & \delta \left( \mathbb{P}(X_0 > 0) + \sum_{i=1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \mathbb{P}(X_i > 0) \right) \ . \end{aligned}$$

Recall that $X_0 = X_1 = \cdots = X_\kappa$, hence

$$\mathbb{P}(X_i > 0) = \mathbb{P}(X_0 > 0) \quad \forall 0 \leq i \leq \kappa \ .$$

For larger values of the index $i$, we note that by the definition of the shrunk graphs, if $W_e^\varphi$ is a positive phase in $G_i$, then $W_e^\varphi \geq (1+\epsilon)^{i-\kappa}$, thus

$$\begin{aligned} \mathbb{P}(X_i > 0) &= \mathbb{P}(X_i \geq (1+\epsilon)^{i-\kappa}) \\ &\leq \mathbb{P}(X_{i-\kappa} \geq (1+\epsilon)^{i-\kappa}) \quad \forall \kappa < i \leq N-1 \ . \end{aligned}$$

Therefore,

$$\begin{aligned}
&|\mathcal{A}' - \mathcal{A}''| \\
&\leq \delta\Bigg(\mathbb{P}(X_0 > 0) + \sum_{i=1}^{\kappa} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \mathbb{P}(X_0 > 0) \\
&\quad + \sum_{i=\kappa+1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \mathbb{P}(X_i > 0)\Bigg) \\
&\leq \delta\Bigg(\mathbb{P}(X_0 > 0)\left(1 + \epsilon\frac{(1+\epsilon)^{\kappa} - 1}{\epsilon}\right) \\
&\quad + \sum_{i=\kappa+1}^{N-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \mathbb{P}(X_{i-\kappa} \geq (1+\epsilon)^{i-\kappa})\Bigg) \\
&= \delta \cdot (1+\epsilon)^{\kappa}\Bigg(\mathbb{P}(X_0 > 0) \\
&\quad + \sum_{i=1}^{N-\kappa-1} \epsilon \cdot (1+\epsilon)^{i-1} \cdot \mathbb{P}(X_i \geq (1+\epsilon^i))\Bigg) \\
&\leq \delta \cdot (1+\epsilon)^{\kappa} \cdot \mathcal{A}' \ .
\end{aligned}$$

By the choice of $\delta = \frac{\epsilon}{(1+\epsilon)^{\kappa}}$, it follows that

$$|\mathcal{A}' - \mathcal{A}''| \leq \epsilon \cdot \mathcal{A}'$$

which, in combination with (3.2) and (3.3), establishes Theorem 3.1.

## 4 Procedure `Estimate`

In this section we present and analyze Procedure `Estimate`. Let $H$ be some $n$-vertex RW graph and consider some positive real $x$ and two performance parameters $\epsilon, \hat{\epsilon} > 0$. Given $H$ and $x$, Procedure `Estimate` runs in time $\mathrm{poly}(n, 1/\epsilon, \log(1/\hat{\epsilon}))$ and outputs an $(\epsilon \cdot \mathbb{P}(\mathcal{X}(H) > 0), \hat{\epsilon})$-approximator of $\mathbb{P}(\mathcal{X}(H) \geq x)$.

In Section 5 we show that the RW graph $H$ can be efficiently transformed into a WNR graph $G$ with the guarantee that the random variable $\mathcal{X}(G)$ is stochastically equivalent to the random variable $\mathcal{X}(H)$. We denote the random variable $\mathcal{X}(G)$ by $X$ and subsequently focus on producing an $(\epsilon \cdot \mathbb{P}(X > 0), \hat{\epsilon})$-approximator of $\mathbb{P}(X \geq x)$.

Fix $E = E(G)$. It will be convenient to partition the edges in $E$ according to their weights to $E^0 = \{e \in E \mid w(e) = 0\}$ and to $E - E^0 = \{e \in E \mid w(e) > 0\}$. By Requirement (R3), we may assume that $p_e > 0$ (and hence $0 < p_e < 1$) for every edge $e = (u, v) \in E^0$ as otherwise, the vertices $u$ and $v$ can be contracted. Let $G^0$ be the restriction of $G$ to the edges in $E^0$.

Let $P_0$ denote the probability of the event $X = \mathcal{X}(G) > 0$. Requirement (R6) implies that this event depends only on the probability space $E^0$; specifically, $\mathcal{X}(G) > 0$ if and only if $\mathcal{X}(G^0) > 0$. Observe that if $G^0$ is disconnected, then $\mathcal{X}(G^0) = \infty$ and the event $\mathcal{X}(G) > 0$ occurs with probability 1. Assuming that $G^0$ is connected, we employ requirement (R6) once more to conclude that $\mathcal{X}(G^0) > 0$ if and only if the edges that fail under the probability space $E^0$ induce a cut on $G^0$.

Fix $c = (5 + \sqrt{17})/2 \approx 4.56$. If $P_0$ is sufficiently large, specifically, at least $n^{-c}$, then the desired approximation can be obtained by a direct Monte Carlo method. Indeed, Theorem 2.1 guarantees that a Monte Carlo method with $O(\log(1/\hat{\epsilon})n^c/\epsilon^2) = \mathrm{poly}(n, 1/\epsilon, \log(1/\hat{\epsilon}))$ trials suffices to generate an $(\epsilon \cdot P_0, \hat{\epsilon})$-approximator of $\mathbb{P}(X \geq x)$. (The random variable for which we apply the Monte Carlo method is simply the indicator of the event $X \geq x$.) This applies in particular to the case where $G^0$ is disconnected which means that $P_0 = 1$. Therefore in what follows we may assume that $P_0 < n^{-c}$ and in particular, that $G^0$ is connected. Note that if $P_0$ is extremely small (e.g., exponentially small in $n$), then the above Monte Carlo method requires too many samples in order to obtain an $(\epsilon \cdot P_0, \hat{\epsilon})$-approximator of $\mathbb{P}(X \geq x)$.

**Dealing with small $P_0$.** How do we efficiently generate an $(\epsilon \cdot P_0, \hat{\epsilon})$-approximator of $\mathbb{P}(X \geq x)$ when $P_0$ is small? For that purpose we introduce the real valued random variable $Y$ which is defined over the probability space $E - E^0$ by mapping each instance $\mathcal{I} : E - E^0 \to \{0, 1\}$ to $\mathbb{P}(X \geq x \mid \mathcal{I})$, namely, $\mathcal{I}$ is mapped to the probability that $\mathcal{X}(G)$ is at least $x$ conditioned on $\mathcal{I}$. This can be viewed as decomposing the probability space $E$ into the Cartesian product of the probability spaces $E - E^0$, from which $\mathcal{I}$ is chosen, and $E^0$, over which $\mathbb{P}(X \geq x \mid \mathcal{I})$ is defined. A crucial observation here is that

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{\mathcal{I}: E - E^0 \to \{0,1\}} \mathbb{P}(\mathcal{I}) \cdot \mathbb{P}(X \geq x \mid \mathcal{I}) \\
&= \sum_{\mathcal{I}: E - E^0 \to \{0,1\}} \mathbb{P}(X \geq x \wedge \mathcal{I}) \\
&= \mathbb{P}(X \geq x) \ ,
\end{aligned}$$

hence our goal is to provide a good approximation for $\mathbb{E}[Y]$. Another important observation is that

$$\mathbb{P}(X \geq x \mid \mathcal{I}) \ \leq \ \mathbb{P}(X > 0 \mid \mathcal{I}) \ = \ \mathbb{P}(X > 0) \ = \ P_0$$

for every instance $\mathcal{I} : E - E^0 \to \{0, 1\}$ (recall that the event $X > 0$ does not depend on the probability space $E - E^0$), therefore $Y \in [0, P_0]$ with probability 1.

Fix $k = 2\ln(4/\hat{\epsilon})/\epsilon^2$ and repeat the following process for $j = 1, \ldots, k$. Choose some instance $\mathcal{I}_j : E - E^0 \to \{0, 1\}$ with probability $\mathbb{P}(\mathcal{I}_j)$ (this can be easily generated by randomly deciding for each edge in $E - E^0$, whether it fails or not independently of all other

edges) and let $Y_j = \mathbb{P}(X \geq x \mid \mathcal{I}_j)$; in other words, $Y_j$ is a random sample of $Y$. Unfortunately, we do not know how to efficiently compute the exact value of $Y_j$ for a given instance $\mathcal{I}_j$. Instead, we will generate an *approximate sample* $Y'_j$ which is an $(\epsilon \cdot P_0/2, \hat{\epsilon}/(2k))$-approximator of $Y_j$.

We will soon explain how the $Y'_j$s are generated, but first let us explain how they are employed to obtain the desired $(\epsilon \cdot P_0, \hat{\epsilon})$-approximator of $\mathbb{E}[Y]$. Let $\bar{Y} = \sum_{j=1}^{k} Y_j/k$ and $\bar{Y}' = \sum_{j=1}^{k} Y'_j/k$. Theorem 2.2 guarantees that $\bar{Y}$ is an $(\epsilon \cdot P_0/2, \hat{\epsilon}/2)$-approximator of $\mathbb{E}[Y]$. By Proposition 2.2, we conclude that $\bar{Y}'$ is an $(\epsilon \cdot P_0/2, \hat{\epsilon}/2)$-approximator of $\bar{Y}$. Therefore Proposition 2.1 implies that $\bar{Y}'$ is an $(\epsilon \cdot P_0, \hat{\epsilon})$-approximator of $\mathbb{E}[Y] = \mathbb{P}(X \geq x)$ as desired.

**Generating the approximate samples.** It remains to present the process through which the approximate samples $Y'_j$ are generated (recall that each approximate sample should be an $(\epsilon \cdot P_0/2, \hat{\epsilon}/(2k))$-approximator of $\mathbb{P}(X \geq x \mid \mathcal{I})$ for some given instance $\mathcal{I} : E - E^0 \to \{0, 1\}$). The technique we use for this process is an extension of Karger's technique [26]. In order to simplify the description of this process, we first assume that there exists some real $p$ such that $p_e = p$ for all edges $e \in E^0$. This assumption is removed later on.

Given some compact cut $C$ of $G^0$, let $\mathcal{F}(C)$ denote the event that all edges in $E(C) \subseteq E^0$ fail. Let $C$ be some min cut of $G^0$ and let $\chi = |E(C)|$ be its size. Since $\mathbb{P}(\mathcal{F}(C)) = p^\chi$, the assumption that $P_0 < n^{-c}$ implies that $p^\chi < n^{-c}$. The following two theorems are established in [26] for the case of 2-way cuts. Building upon the recent bound of Berend & Tassa on the Bell number [5], we extend them to (compact) $r$-way cuts for all $r \geq 2$ simultaneously.

THEOREM 4.1. *For every real $\alpha \geq 1$, there are less than $13n^{2\alpha}$ compact cuts of size at most $\alpha\chi$ in $G^0$.*

*Proof.* The theorem is established by presenting a random process that generates each compact cut of size at most $\alpha\chi$ in $G^0$ with probability greater than $\frac{1}{13}n^{-2\alpha}$. Observe first that if an $r$-way cut $C$ satisfies $|E(C)| \leq \alpha\chi$, then $r$ must be at most $2\alpha$ as otherwise there exists some cluster $U$ of $C$ with less than $\chi$ edges crossing between $U$ and $V(G^0) - U$, in contradiction to the assumption that $\chi$ is the size of a min cut of $G^0$.

Fix $k = \lceil 2\alpha \rceil$. Our random process first performs random edge contractions in $G^0$ until $k$ vertices $v_1, \ldots, v_k$ remain in the graph (cf. Section 2.2.1 in [26]); each vertex $v_i$ corresponds to some subset $V_i \subseteq V(G^0)$ so that $\{V_1, \ldots, V_k\}$ is a partition of $V(G^0)$ (the subgraph induced on $G^0$ by $V_i$ is connected). We then take $P = \{Q_1, \ldots, Q_\ell\}$, $1 \leq \ell \leq k$, to be a partition of $\{v_1, \ldots, v_k\}$ chosen uniformly at random out of the $B_k$ possible partitions of $\{v_1, \ldots, v_k\}$, where $B_k$ is the $k^{\text{th}}$ *Bell number*. The cut $\widehat{C} = \{U_1, \ldots, U_\ell\}$ generated by our random process is defined by setting $U_j = \bigcup\{V_i \mid v_i \in Q_j\}$ for $j = 1, \ldots, \ell$. It is important to note that $\widehat{C}$ is not necessarily a compact cut, however, if $C$ is any compact cut of size at most $\alpha\chi$ in $G^0$, then it can be generated by our random process. Our goal in the remainder of this proof is to show that $C$ is indeed generated with probability greater than $\frac{1}{13}n^{-2\alpha}$.

Karger [26] shows that the probability that none of the edges crossing $C$ is contracted during the random edge contractions is at least

$$\left(1 - \frac{2\alpha}{n}\right)\left(1 - \frac{2\alpha}{n-1}\right)\cdots\left(1 - \frac{2\alpha}{k+1}\right) = \frac{\binom{k}{2\alpha}}{\binom{n}{2\alpha}} ,$$

where generalized binomial coefficients[7] are used when $2\alpha$ is not an integer. It remains to prove that $\frac{\binom{k}{2\alpha}}{\binom{n}{2\alpha}} B_k^{-1} > \frac{1}{13}n^{-2\alpha}$. Indeed,

$$\begin{aligned} \frac{\binom{k}{2\alpha}}{\binom{n}{2\alpha}} B_k^{-1} &\geq \left(\frac{k}{2\alpha}\right)^{2\alpha}\left(\frac{2\alpha}{en}\right)^{2\alpha} B_k^{-1} \\ &> \left(\frac{k}{en}\right)^{2\alpha}\left(\frac{\ln(k+1)}{0.792k}\right)^{k} \qquad (4.4) \\ &> n^{-2\alpha} \cdot \frac{1}{k}\left(\frac{\ln(k+1)}{0.792e}\right)^{k} , \end{aligned}$$

where inequality 4.4 is due to Berend & Tassa [5]. The assertion follows as $\frac{1}{k}\left(\frac{\ln(k+1)}{0.792e}\right)^{k} > \frac{1}{13}$ when $k \geq 2$.

THEOREM 4.2. *For every real $\alpha \geq 1$, the probability that there exists some compact cut $C$ of size at least $\alpha\chi$ in $G^0$ such that $\mathcal{F}(C)$ occurs is $O(n^{-\alpha\eta})$, where $\eta$ is defined by fixing $p^\chi = n^{-(2+\eta)}$.*

*Proof.* Let $C_1, \ldots, C_t$ be the compact cuts of size at least $\alpha\chi$ and for each $1 \leq i \leq t$, let $\chi_i = |E(C_i)|$. We assume without loss of generality that $\alpha\chi \leq \chi_1 \leq \cdots \leq \chi_t$. Denote $p_i = p^{\chi_i} = \mathbb{P}(\mathcal{F}(C_i))$ and consider some real $\beta \geq 1$. By Theorem 4.1, there are less than $13n^{2\beta}$ compact cuts of size at most $\beta\chi$. It follows that $\chi_{13n^{2\beta}}$ must be greater than $\beta\chi$.

[7] Generalized binomial coefficients are a generalization of the standard binomial coefficients $\binom{x}{y}$ to non-integral $x$ and $y$. This generalization is based on replacing the factorial in the standard definition with the Gamma function. Many of the identities and bounds that hold for the standard binomial coefficients also hold in the generalized case, including the bounds $\left(\frac{x}{y}\right)^y \leq \binom{x}{y} \leq \left(\frac{ex}{y}\right)^y$.

We shall bound the probability that $\mathcal{F}(C_i)$ occurs for some (at least one) $1 \leq i \leq t$ by bounding the sum $\sum_{i=1}^{t} p_i$. The first $t' = 13n^{2\alpha+1/\ln(n)}$ terms are bounded simply by observing that

$$\begin{aligned}\sum_{i=1}^{t'} p_i &\leq 13n^{2\alpha+1/\ln(n)} \cdot p^{\alpha\chi} \\ &= 13e \cdot n^{2\alpha} \cdot n^{-\alpha(2+\eta)} = 13e \cdot n^{-\alpha\eta} \ .\end{aligned}$$

Thus it remains to bound the remaining $t - t'$ terms.

Given some $\beta \geq \alpha$, we write $s = 13n^{2\beta+1/\ln(n)}$ and conclude that $\chi_s > \left(\beta + \frac{1}{2\ln(n)}\right)\chi = \frac{\ln(s)-\ln(13)}{2\ln n} \cdot \chi$. Therefore

$$\begin{aligned}p_s &< (p^{\chi})^{\frac{\ln(s)-\ln(13)}{2\ln n}} \\ &= \left(n^{-(2+\eta)}\right)^{\frac{\ln(s)-\ln(13)}{2\ln n}} \\ &= \left(e^{\frac{\ln(s)-\ln(13)}{2}}\right)^{-(2+\eta)} \\ &= s^{-(1+\eta/2)} \cdot 13^{1+\eta/2} \ .\end{aligned}$$

Summing over all $i > t'$, we get

$$\begin{aligned}\sum_{i>t'} p_i &< 13^{1+\eta/2} \cdot \sum_{s>t'} s^{-(1+\eta/2)} \\ &\leq 13^{1+\eta/2} \cdot \int_{t'}^{\infty} s^{-(1+\eta/2)} ds \\ &= 13^{1+\eta/2} \cdot \left(-(\eta/2) \cdot s^{-\eta/2}\Big|_{13n^{2\alpha+1/\ln(n)}}^{\infty}\right) \\ &= 13^{1+\eta/2} \cdot (\eta/2) \cdot 13^{-(\eta/2)} \cdot n^{-\alpha\eta} \cdot e^{-\eta/2} \\ &\leq 13 \cdot n^{-\alpha\eta} \ .\end{aligned}$$

The assertion follows.

Notice that the compact cuts addressed in Theorems 4.1 and 4.2 may have an arbitrary number of clusters, but their size is compared to $\alpha\chi$, where $\chi$ is the size of the smallest 2-way cut in $G^0$. This point is crucial for the validity of the arguments.

Write $p^{\chi} = n^{-(2+\eta)}$. We must have $\eta > c - 2$ as $p^{\chi} < n^{-c}$. Fix $\alpha = \frac{c-1+\ln(1/\epsilon)/\ln(n)}{2}$ and let $\mathcal{C}$ be the collection of all compact cuts of size at most $\alpha\chi$ in $G^0$. By Theorem 4.1, $\mathcal{C}$ consists of $O(n^{2\alpha}) = O(n^{c-1}/\epsilon)$ compact cuts. These cuts can be enumerated with high probability in time $\tilde{O}(n^{2\alpha})$ by the algorithm of [28]; they can also be enumerated deterministically by the slightly slower algorithm of [40].

Given some collection $\mathcal{B}$ of compact cuts in $G^0$, let

$$\psi(\mathcal{B}) = \mathbb{P}\left(\bigvee_{C \in \mathcal{B}} \mathcal{F}(C)\right)$$

be the probability that all crossing edges of some (at least one) cut in $\mathcal{B}$ fail. Theorem 4.2 guarantees that $0 \leq P_0 - \psi(\mathcal{C}) \leq \gamma n^{-\alpha\eta}$ for some universal constant $\gamma$. The choice of $c = (5 + \sqrt{17})/2$ and of $\alpha = \frac{c-1+\ln(1/\epsilon)/\ln(n)}{2}$ and the assumption that $\eta > c-2$ ensure that $\gamma n^{-\alpha\eta} \leq \epsilon n^{-(2+\eta)}/4 = \epsilon p^{\chi}/4$ as long as $(4\gamma)^4 \leq (1/\epsilon)^{c-4}$, which yields the following corollary.

COROLLARY 4.1. *The probability that there exists some compact cut $C \notin \mathcal{C}$ such that $\mathcal{F}(C)$ occurs is at most $\epsilon p^{\chi}/4 \leq \epsilon \cdot P_0/4$.*

Consider some instance $\mathcal{I} : E - E^0 \to \{0,1\}$. Our goal is to efficiently generate an $(\epsilon \cdot P_0/2, \hat{\epsilon}/(2k))$-approximator of $\mathbb{P}(X \geq x \mid \mathcal{I})$. For a given compact $r$-way cut $C = \{U_1, \ldots, U_r\}$ of $G^0$, we construct the graph $G_{C,\mathcal{I}}$ as follows. The vertex set of $G_{C,\mathcal{I}}$ is $V(G_{C,\mathcal{I}}) = \{u_1, \ldots, u_r\}$. For every edge $e \in E - E^0$ with one endpoint in the cluster $U_i$ and the other in the cluster $U_j$, $i \neq j$, such that $\mathcal{I}(e) = 1$ (that is, $e$ does not fail under $\mathcal{I}$), we add an edge $(u_i, u_j)$ to $E(G_{C,\mathcal{I}})$ whose weight is $w(e)$. The following proposition is due to requirement (R3).

PROPOSITION 4.1. *Conditioned on the instance $\mathcal{I} : E - E^0 \to \{0,1\}$, and on the event that the set of failing edges in $E_0$ induces the compact $r$-way cut $C$ on $G^0$, we have $\mathcal{X}(G_{C,\mathcal{I}}) = X$.*

Let $\mathcal{B}_{\mathcal{I}}$ be the collection of all compact cuts $C$ of $G^0$ such that $\mathcal{X}(G_{C,\mathcal{I}}) \geq x$. Proposition 4.1 implies that $\mathbb{P}(X \geq x \mid \mathcal{I}) = \psi(\mathcal{B}_{\mathcal{I}})$. By Corollary 4.1, we know that $\psi(\mathcal{B}_{\mathcal{I}} - \mathcal{C}) \leq \epsilon \cdot P_0/4$, and hence $\psi(\mathcal{B}_{\mathcal{I}}) - \epsilon \cdot P_0/4 \leq \psi(\mathcal{B}_{\mathcal{I}} \cap \mathcal{C}) \leq \psi(\mathcal{B}_{\mathcal{I}})$. Consequently, it suffices to generate an $(\epsilon \cdot P_0/4, \hat{\epsilon}/(2k))$-approximator of $\psi(\mathcal{B}_{\mathcal{I}} \cap \mathcal{C})$.

**Probabilistic DNF satisfiability.** The approximation of $\psi(\mathcal{B}_{\mathcal{I}} \cap \mathcal{C})$ is performed by the method of Karp, Luby, and Madras [29, 30] for approximating the probability that a formula in disjunctive normal form (DNF) is satisfied. Given some DNF formula $\phi$, and given the probability $q_i$ that $x_i$ is assigned to true for each variable $x_i$ (independently of all other variables), the method of Karp et al. generates a $(\delta \cdot q(\phi), \hat{\delta})$-approximator of the probability $q(\phi)$ that $\phi$ is satisfied in time $O(|\phi| \log(1/\hat{\delta})/\delta^2)$, where $|\phi|$ stands for the size of the formula (number of literals).

Cast into that framework, the event $\bigvee_{C \in \mathcal{B}_{\mathcal{I}} \cap \mathcal{C}} \mathcal{F}(C)$ is encoded as a DNF formula whose variables correspond to whether or not the edges in $E^0$ fail and whose clauses correspond to the cuts in $\mathcal{B}_{\mathcal{I}} \cap \mathcal{C}$. Such a DNF formula has $|\mathcal{B}_{\mathcal{I}} \cap \mathcal{C}| \leq |\mathcal{C}| = O(n^{c-1}/\epsilon)$ clauses, each with at most $n$ literals. Therefore an $(\epsilon \cdot \psi(\mathcal{B}_{\mathcal{I}} \cap \mathcal{C})/4, \hat{\epsilon}/(2k))$-approximator of $\psi(\mathcal{B}_{\mathcal{I}} \cap \mathcal{C})$, which also serves as an $(\epsilon \cdot P_0/4, \hat{\epsilon}/(2k))$-approximator of $\psi(\mathcal{B}_{\mathcal{I}} \cap \mathcal{C})$ since $\psi(\mathcal{B}_{\mathcal{I}} \cap$

$\mathcal{C}) \leq P_0$, can be generated in time $O(\log(k/\hat{\epsilon})n^c/\epsilon^3) = \text{poly}(n, 1/\epsilon, \log(1/\hat{\epsilon}))$.

**Varying failure probabilities.** Recall that in attempt to simplify the description of the process that generates approximate samples of the random variable $Y$, we assumed that $p_e = p$ for all edges $e \in E^0$. We now turn to lift this assumption. The technique we use here is essentially identical to that used by Karger [26] for a similar purpose; we describe it for completeness.

The WNR graph $G^0$ with varying failure probabilities $0 < p_e < 1$ is transformed into a WNR graph $H$, $V(H) = V(G^0)$, with uniform failure probabilities $p = 1 - \theta$ for some sufficiently small $\theta > 0$. For each edge $e = (u, v) \in E^0$ with failure probability $0 < p_e < 1$, we introduce a *bundle* of $k_e = \ln(1/p_e)/\theta$ parallel $(u, v)$ edges in $H$; all edges in $H$ have zero weight (just like all edges in $G^0$ have zero weight). The probability that all $k_e$ edges in this bundle fail is $(1-\theta)^{\ln(1/p_e)/\theta}$ which converges to $p_e$ as $\theta \to 0$. By requirement (R3), it is sufficient to generate approximate samples for the random variable $Y$ with respect to the graph $H$ in the limit as $\theta \to 0$. The technique we introduced earlier in this section is suitable for that as $H$ has uniform failure probabilities. In particular it is sufficient to enumerate all the small compact cuts $C$ of $H$, identify those inducing $\mathcal{X}(G_{C,\mathcal{I}}) \geq x$ for a given instance $\mathcal{I} : E - E^0 \to \{0, 1\}$, and then approximate the probability that all crossing edges of at least one of them fail.

Note that changing the parameter $\theta$ scales the size of cuts in $H$ without changing their relative sizes. We construct a graph $H'$, $V(H') = V(G^0)$, with positive costs on the edges by assigning cost $\ln(1/p_e)$ to each edge $e \in E^0$. The small cost cuts in $H'$ correspond to the small sized cuts in $H$; they can be enumerated by known techniques.

Given the small cuts in $H$ that induce $\mathcal{X}(G_{C,\mathcal{I}}) \geq x$, we have to approximate the probability, as $\theta \to 0$, that all crossing edges of at least one of them fail. We already argued that this is exactly the probability that all crossing edges of at least one of the corresponding cuts in $G^0$ fail. Approximating this probability is done as before by constructing the appropriate DNF formula and employing the method of Karp et al. [29, 30].

## 5 Transforming RW graphs into WNR graphs

In this section we present an efficient transformation that takes an RW graph $G$ and outputs a WNR graph $G'$ such that the random variable $\mathcal{X}(G')$ is stochastically equivalent to the random variable $\mathcal{X}(H)$. This is similar to a method presented by Mirchandani [33] (see also [3]).

Let $G$ be an arbitrary RW graph and consider some edge $e \in E(G)$. Recall that there exist some positive integer $m(e)$ and some non-negative phases $W_e^1, \ldots, W_e^{m(e)}$ and probabilities $p_e^1, \ldots, p_e^{m(e)}$, where $\sum_{i=1}^{m(e)} p_e^i = 1$, such that $w(e) = W_e^i$ with probability $p_e^i$ independently of all other edges. We assume without loss of generality that the phases of $e$ are distinct (identical phases can be merged into one) and ordered so that $W_e^1 < \cdots < W_e^{m(e)}$.

The WNR graph $G'$ is obtained by taking $V(G') = V(G)$ and transforming every edge $e = (u, v) \in E(G)$ into $m(e)$ parallel $(u, v)$-edges $e_1, \ldots, e_{m(e)}$ in $G'$. The weight of edge $e_i$ is set to $w(e_i) \leftarrow W_e^i$ for each $1 \leq i \leq m(e)$. The failure probabilities $p_{e_1}, \ldots, p_{e_{m(e)}}$ are designed to guarantee that the random variable $M_e = \min\{w(e_i) \mid 1 \leq i \leq m(e), e_i \text{ did not fail}\}$ in $G'$ is stochastically equivalent to the random variable $w(e)$ in $G$. This is achieved by setting

$$p_{e_i} \leftarrow 1 - \frac{p_e^i}{1 - \sum_{j=1}^{i-1} p_e^j} = \frac{1 - \sum_{j=1}^{i} p_e^j}{1 - \sum_{j=1}^{i-1} p_e^j}$$

for every $1 \leq i \leq m(e) - 1$; and $p_{e_{m(e)}} \leftarrow 0$. Indeed, for every $1 \leq i \leq m(e)$, we have

$$\begin{aligned}
\mathbb{P}(M_e = W_e^i) &= (1 - p(e_i)) \cdot \prod_{j=1}^{i-1} p(e_j) \\
&= \frac{p_e^i}{1 - \sum_{j=1}^{i-1} p_e^j} \cdot \prod_{j=1}^{i-1} \frac{1 - \sum_{l=1}^{j} p_e^l}{1 - \sum_{l=1}^{j-1} p_e^l} \\
&= p_e^i \ ,
\end{aligned}$$

where the last equation holds by telescoping. Requirement (R4) implies that $\mathcal{X}(G')$ is stochastically equivalent to $\mathcal{X}(G)$.

## 6 Hardness

In this section we prove that the problem of computing the expected diameter of an RW graph is #P-hard. The problem remains #P-hard even when restricted to identically distributed weighted graphs. Our line of arguments immediately implies that computing the radius of an identically distributed weighted graph with respect to a designated vertex is also #P-hard.

Hardness is established by reduction from the TTNR problem defined as follows. On input connected graph $G$ and two vertices $s, t \in V(G)$, the goal is to compute the probability $P_G$ that $s$ and $t$ remain in the same connected component when each edge in $E(G)$ fails with probability $1/2$ independently of all other edges. The #P-hardness of TTNR is established by Valiant [39]. Since the support of $P_G$ consists of integer multiples of $2^{-m}$, where $m = |E(G)|$, we conclude that it is #P-hard to approximate $P_G$ to within a one-sided additive error of $\epsilon$ for any $\epsilon < 2^{-m}$.

Given a graph $G$ with two vertices $s, t \in V(G)$ as input of TTNR, we construct an identically distributed weighted graph $G'$ with parameters $m(e) = 2$, $p_e^1 = p_e^2 = 1/2$, $W_e^1 = 0$, and $W_e^2 = 1$ for every edge $e \in E(G')$. $G'$ is obtained from $G$ by adding a new edge $e = (s, t)$ and augmenting the resulting graph with two simple paths, one connecting $s$ to the new vertex $s'$ and the other connecting $t$ to the new vertex $t'$. Each new simple path consists of $k = \Theta(m)$ new vertices. The reduction is cast in the following lemma.

LEMMA 6.1. *Let $D$ be the random variable that takes on* $\mathrm{diam}(G')$. *Then*

$$P_G - 2^{-m} < 1 - 2(\mathbb{E}[D] - k) \leq P_G \ .$$

*Proof.* Let $D'$ be the random variable that takes on $\mathrm{dist}_{G'}(s', t')$. By definition, we know that $D' \leq D$ with probability 1. We shall take $k$ to be sufficiently large so that Chernoff's inequality implies that $\mathbb{P}(\mathrm{dist}_{G'}(s', s) < n \vee \mathrm{dist}_{G'}(t, t') < n) < 2^{-(m+1)}/n$, where $n = |V(G)|$. By the construction of $G'$, it follows that $\mathbb{P}(D > D') < 2^{-(m+1)}/n$. Since $D - D' < n$, we conclude that $0 \leq \mathbb{E}[D] - \mathbb{E}[D'] < 2^{-(m+1)}$.

By the linearity of expectation, we have

$$\begin{aligned}\mathbb{E}[D'] &= \mathbb{E}[\mathrm{dist}_{G'}(s', s)] + \mathbb{E}[\mathrm{dist}(s, t)] + \mathbb{E}[\mathrm{dist}_{G'}(t, t')] \\ &= k + \mathbb{E}[\mathrm{dist}(s, t)] \ ,\end{aligned}$$

where the last term can be rewritten as

$$\begin{aligned}\mathbb{E}[\mathrm{dist}(s, t)] &= \mathbb{E}[\mathrm{dist}(s, t) \mid w(e) = 1] \cdot \mathbb{P}(w(e) = 1) \\ &\quad + \mathbb{E}[\mathrm{dist}(s, t) \mid w(e) = 0] \cdot \mathbb{P}(w(e) = 0) \\ &= \mathbb{E}[\mathrm{dist}(s, t) \mid w(e) = 1]/2 \ .\end{aligned}$$

The assertion is established by arguing that $\mathbb{E}[\mathrm{dist}(s, t) \mid w(e) = 1] = 1 - P_G$. Indeed, when $w(e) = 1$, then $\mathrm{dist}_{G'}(s, t) \in \{0, 1\}$. The argument holds since the $\mathrm{dist}_{G'}(s, t) = 0$ instances (respectively, the $\mathrm{dist}_{G'}(s, t) = 1$ instances) of the probability space $E(G')$ correspond to the instances of TTNR in which $s$ and $t$ remain connected (resp., become disconnected).

## 7 Conclusions

We study the setting of graphs whose edge weights are independent random variables and show that for the wide family of efficiently calculated[8] distance-cumulative weighted graph properties, the problem of computing the $k^{\text{th}}$ moment admits an FPRAS. This turns out to be a non-trivial task; in particular, a straightforward application of the Monte Carlo method for approximating, e.g., the expectation (i.e., the first moment) fails when the variance is large as too many samples are required in order to take into account low probability events[9] that may drastically affect the expectation.

Our technique does not guarantee a (multiplicative) approximation for the $k^{th}$ *central moment* (and in particular, the variance) when this is close to zero. (This is in contrast to the $k^{\text{th}}$ moment *about zero* for which the approximation is guaranteed regardless.) However, it does provide us with the ability to decide if the $k^{\text{th}}$ central moment is indeed close to zero.

There are still some fundamental weighted graph properties which are not distance-cumulative, and hence cannot be dealt with via our technique, such as the shortest $(s, t)$-path and the weight of a maximum matching. It is also natural to consider the directed analogue of randomly weighted graphs and in particular, various network flow problems. Another aspect that calls for further research, once encoding issues are resolved, is that of continuous distributions for the edge weights.

**Acknowledgment**

We would like to thank Noga Alon for helpful discussions and the anonymous reviewers for carefully reading the paper and suggesting valuable remarks.

# APPENDIX

## A High critical ratio

Consider the identically distributed weighted graph $G$ consisting of 2 vertices and $m$ parallel edges connecting them, where each edge is of weight 1 with probability 1/2; and of weight $2^{2m}$ otherwise. Let $X$ denote the random variable that takes on the diameter of $G$. It is easy to verify that $\mathbb{E}[X] \approx 2^m$, while $\mathrm{Var}[X] \approx 2^{3m}$, so the critical ratio here is roughly $2^m$. Indeed, a Monte Carlo method with significantly less than $2^m$ samples would most probably estimate the expected diameter of $G$ to be 1 which is an awful approximation.

## References

[1] C. Alexopoulos and J.A. Jacobson. State space partition algorithms for stochastic systems with applications

[8] In fact, our results also hold for distance-cumulative weighted graph properties that can be efficiently approximated to within a multiplicative error of $\rho > 1$. For such properties our scheme only guarantees a $\rho(1 + \epsilon)$-approximation.

[9] Events occurring with probability which is low with respect to the limited computational resources of our algorithm are still expected to occur in practice if the network in hand (say, the Internet) is being used very frequently.

to minimum spanning trees. *Networks*, 35(2):118–138, 2000.

[2] N. Alon, A. Frieze, and D. Welsh. Polynomial time randomized approximation schemes for Tutte-Gröthendieck invariants: the dense case. *Random Struct. Algorithms*, 6(4):459–478, 1995.

[3] M.O. Ball, C.J. Colbourn, and J.S. Provan. Network reliability. *Handbook of Operations Research: Network Models*, pp. 673–762, Elsevier North-Holland, 1995.

[4] N. Banerjee, M.D. Corner, and B.N. Levine. An energy-efficient architecture for DTN throwboxes. In *Proc. IEEE Infocom*, 2007.

[5] D. Berend and T. Tassa. Improved bounds on Bell numbers and on moments of sums of random variables. To appear in *Probability and Mathematical Statistics*.

[6] D.J. Bertsimas and G. van Ryzin. An asymptotic determination of the minimum spanning tree and minimum matching constants in geometrical probability. *Oper. Res. Lett.*, 9:223–231, 1990.

[7] A. Beveridge, A.Frieze, and C. McDiarmid. Random minimum length spanning trees in regular graphs. *Combinatorica*, 18:311–333, 1998.

[8] B. Bollobás. *Modern Graph Theory*. Graduate texts in mathematics, Springer-Verlag, Berlin, New York, 1998.

[9] S. Burleigh, A. Hooke, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, and H. Weiss. Delay-tolerant networking: an approach to interplanetary internet. *IEEE Communications Magazine*, 41(6):128–136, 2003.

[10] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 23(4):493–507, 1952.

[11] M. Coates, A. Hero, R. Nowak, and B. Yu. Internet tomography. *IEEE Signal Processing Magazine*, 19:47–65, 2002.

[12] C. Cooper, A.M. Frieze, K. Mehlhorn, and V. Priebe. Average-case complexity of shortest-paths problems in the vertex-potential model. *Random Struct. Algorithms*, 16(1):33–46, 2000.

[13] N.G. Duffield and F.Lo Presti. Multicast inference of packet delay variance at interior network links. In *Proc. IEEE INFOCOM*, 2000.

[14] S. Even. *Graph Algorithms*. Computer Science Press, 1979.

[15] K. Fall. A delay-tolerant network architecture for challenged internets. In *Proc. ACM SIGCOMM*, pages 27–34, 2003.

[16] A.M. Frieze. On the value of a random minimum spanning tree problem. *Discrete Appl. Math.*, 10:47–56, 1985.

[17] A.M. Frieze and G.R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Appl. Math.*, 10:57–77, 1985.

[18] A.M. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Struct. Algorithms*, 10(1-2):5–42, 1997.

[19] D.R. Fulkerson. Expected critical path lengths in PERT Networks. *Oper. Res.*, 10:808–817, 1962.

[20] L.A. Goldberg and M. Jerrum. Inapproximability of the Tutte polynomial. *Inf. Comput.*, 206(7):908–929, 2008.

[21] R. Hassin and A. Tamir. On the minimum diameter spanning tree problem. *Inf. Process. Lett.*, 53(2):109–111, 1995.

[22] R. Hassin and E. Zemel. On shortest paths in graphs with random weights. *Mathematics of Operations Research*, 10(4):557–564, 1985.

[23] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. American Statistical Association*, 58(301):13–30, 1963.

[24] A. Jain and J.W. Mamer. Approximations for the random minimal spanning tree with applications to network provisioning. *Oper. Res.*, 36:575–584, 1988.

[25] S. Janson. One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8:347–361, 1999.

[26] D.R. Karger. A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem. *SIAM J. Comput. (SICOMP)*, 29(2):492–514, 1999.

[27] D.R. Karger, D. Koller, and S.J. Phillips. Finding the hidden path: time bounds for all-pairs shortest paths. *SIAM J. Comput. (SICOMP)*, 22(6):1199–1217, 1993.

[28] D.R. Karger and C. Stein. A new approach to the minimum cut problem. *J. ACM*, 43(4):601–640, 1996.

[29] R.M. Karp and M. Luby. Monte-Carlo algorithms for the planar multiterminal network reliability problem. *J. Complexity (JC)*, 1(1):45–64, 1985.

[30] R.M. Karp, M. Luby, and N. Madras. Monte-Carlo approximation algorithms for enumeration problems. *J. Algorithms*, 10(3):429–448, 1989.

[31] V.G. Kulkarni. Minimal spanning trees in undirected networks with exponentially distributed arc weights. *Networks*, 18(2):111–124, 1988.

[32] U. Mayer. Average-case complexity of single-source shortest-paths algorithms: lower and upper bounds. *J. Algorithms*, 48(1):91–134, 2003.

[33] P.B. Mirchandani. Shortest distance and reliability of probabilistic networks. *Comp. Oper. Res.*, 3:347–355, 1976.

[34] E.F. Moore and C.E. Shannon. Reliable circuits using less reliable relays. *J. Franklin Inst.*, 262:191–208; 263:281–297, 1956.

[35] J. Partan, J. Kurose, and B.N. Levine. A survey of practical issues in underwater networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 11(4), 2007.

[36] Y. Schwartz, Y. Shavitt, and U. Weinsberg. A measurement study of the origins of end-to-end delay variations. In *Proc. Passive and Active Measurement Conference (PAM)*, 2010.

[37] T.L. Snyder and J.M. Steele. Probabilistic networks and network algorithms. *Handbook of Operations Research: Network Models*, pp. 401–424, Elsevier North-Holland, 1995.

[38] J.M. Steele. On Frieze's $\zeta(3)$ limit for lengths of min-

imal spanning trees. *Discrete Appl. Math.*, 18:99–103, 1987.

[39] L.G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comput. (SICOMP)*, 8(3):410–421, 1979.

[40] V.V. Vazirani and M. Yannakakis. Suboptimal cuts: their enumeration, weight and number. In *Proc. 19th International Colloq. on Automata, Languages and Programming (ICALP)*, pp. 366–377, 1992.

[41] X. Zhang, J. Kurose,B.N. Levine, D. Towsley, and H. Zhang. Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing. In *Proc. ACM MobiCom*, 2007.