

Approximating two-terminal network reliability

Weiming Feng*

Yucheng Fu*

Heng Guo†

Abstract

We present a fully polynomial-time randomised approximation scheme (FPRAS) for the two-terminal reliability problem on general graphs, both directed and undirected. We also show that the complementary unreliability question is $\#BIS$ -hard. The key idea of the algorithm was discovered by GPT-5.6 Sol Ultra.

1 Introduction

In 1979, Valiant [Val79] introduced the computational complexity class $\#P$, a counting analogue of NP . One of his main motivations was to capture the intrinsic hardness of computing network reliability measures. Here, the setting is that we are given a stochastic network, which can be a directed or undirected graph. Each arc in a directed graph, or edge in an undirected graph, opens or closes independently with some probability, and reliability is measured by the probability of being connected in the open subgraph. One of the most prominent variant is the two-terminal reliability, which asks for the probability that if a given terminal s can reach another given terminal t in the stochastic graph. Indeed, this is listed as one of the first thirteen $\#P$ -complete problems by Valiant [Val79].

Reliability measures have been extensively studied, and it turns out that most variants are $\#P$ -complete [Bal80, Jer81, BP83, PB83, Bal86, Col87], which means that efficient exact computation is unlikely. For two terminal reliability, it remains $\#P$ -complete when every failure probability is $1/2$ [Val79], on directed acyclic graphs [PB83], or on planar graphs with maximum degree three [Pro86].

Due to the apparent intractability, it is then natural to ask for approximation. However, progress about the approximation complexity of reliability measures is much slower. For these problems, the appropriate notion of approximation is fully polynomial-time randomised approximation schemes (FPRAS), which asks for a relative approximation that runs in polynomial time in the input size and inverse relative error. The first FPRAS in this context is due to Karger [Kar01] for all-terminal undirected unreliability, whereas the complementary all-terminal reliability approximation is solved much later by Guo and Jerrum [GJ19].¹ For two-terminal reliability, much less is known. The special case restricted to directed acyclic graphs (DAGs) admits an FPRAS. This is because the special case puts the problem into a certain complexity class called $SpanL$, and Arenas, Croquevielle, Jayaram, and Riveros [ACJR21] showed that all problems in $SpanL$ admit FPRASes. See also [AvBGM25, FG24].

*School of Computing and Data Science, The University of Hong Kong, HK, China

†School of Informatics, University of Edinburgh, Informatics Forum, Edinburgh, EH8 9AB, UK

E-mail address: wfeng@hku.hk, fyc0130@connect.hku.hk, and hguo@inf.ed.ac.uk.

¹While reliability and unreliability sum to 1, either quantity can be exponentially small, which prevents a relative approximation of one to translate to that of the other.

The approximation complexity of two-terminal reliability in general, despite its old history [KL85, ZL11], remains unresolved. It was explicitly asked in the survey by Kannan [Kan94] if an FPRAS exists. In this paper we answer this question positively, by giving algorithms for both the directed and the undirected case. In fact, we first give an algorithm for the directed version, and then reduce the undirected version to it. The reduction is essentially due to Jerrum [Jer81].

Thus, we formally introduce the directed version next. Let $G = (V, E)$ be a directed graph with two distinguished vertices s, t . Each arc $e \in E$ is associated with an *open probability* $p_e \in (0, 1]$. The *two-terminal network reliability* of G asks the following question: suppose that each arc is independently open with probability p_e , what is the probability that there is a path from s to t ? Equivalently, the probability can be written as

$$\text{Rel}(s, t) = \Pr_{G_p}[s \rightarrow t] = \sum_{\substack{F \subseteq E: \\ s \text{ can reach } t \text{ in } (V, F)}} \prod_{e \in F} p_e \prod_{e \in E \setminus F} (1 - p_e),$$

where G_p denotes the random subgraph of open arcs. In the special case when all $p_e = 1/2$, the problem amounts to counting the spanning subgraphs of G in which s can reach t .

By *FPRAS* (fully polynomial randomised approximation scheme), we mean a randomised algorithm that, given a directed graph $G = (V, E)$, two terminals $s, t \in V$, open probabilities $p_e \in (0, 1]$ for each arc $e \in E$, and an error parameter $\varepsilon > 0$, outputs a random variable $\hat{R}$ such that

$$\Pr \left[(1 - \varepsilon) \text{Rel}(s, t) \leq \hat{R} \leq (1 + \varepsilon) \text{Rel}(s, t) \right] \geq \frac{3}{4},$$

and whose running time is polynomial in the input length and in ε^{-1} .

Theorem 1. *There is an FPRAS for two-terminal reliability on directed graphs. On an input with $n = |V|$ vertices, $m = |E|$ arcs, minimum open probability $p_{\min} = \min_{e \in E} p_e$, its running time is $\tilde{O} \left(n^3 m^2 \left(\frac{1}{\varepsilon^2} + n \log \left(\frac{2}{p_{\min}} \right) \right) \right)$, where $\tilde{O}(\cdot)$ hides factors polylogarithmic in n , ε^{-1} , and $\log \left(\frac{2}{p_{\min}} \right)$.*

The algorithm in Theorem 1 is a combination of simulated annealing and Markov chain Monte Carlo. It is a significant departure from previously known reliability algorithms, such as [Kar01, GJ19, CGZZ24, ACJR21]. In spirit, the closest algorithm is perhaps the celebrated approximation algorithm for non-negative permanents by Jerrum, Sinclair, and Vigoda [JSV04]. We give a high-level overview of the algorithm in Section 1.1.

Another variant of the problem is the *undirected* version, where G is undirected and we ask the probability that s and t are connected in the random subgraph. For this variant, efficient algorithms are known only for graphs of low tree-width [SW85, GM20], and no efficient approximation algorithm is known. Jerrum [Jer81] gave an exact constant-size gadget reduction to the directed version that preserves $p_{\min}$: replace each undirected edge with the directed gadget below.

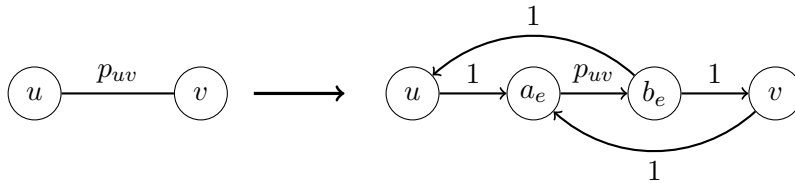


Figure 1: Directed gadget for undirected edge $\{u, v\}$ with open probability p_{uv} .

For a connected undirected graph with n vertices and m edges with $m \geq n - 1$, the reduction produces a directed graph with $n + 2m$ vertices and $5m$ arcs, while preserving the minimum open probability. Combined with [Theorem 1](#), this gives the following result.

Corollary 2. *There is an FPRAS for two-terminal reliability on undirected graphs. On an input with n vertices, m edges, minimum open probability $p_{\min}$, and error parameter $\varepsilon > 0$, its running time is $\tilde{O}\left(m^5 \left(\frac{1}{\varepsilon^2} + m \log\left(\frac{2}{p_{\min}}\right)\right)\right)$.*

In light of Karger’s celebrated algorithm [\[Kar01\]](#), it is also interesting to ask the approximation complexity for unreliability $\text{UnRel}(s, t) = 1 - \text{Rel}(s, t)$. Feng and Guo [\[FG24\]](#) showed that even in DAGs, approximating $s - t$ unreliability is $\#\text{BIS}$ -hard. Here, $\#\text{BIS}$ is the problem of counting the number of independent sets in bipartite graphs. While $\#\text{BIS}$ is not known to be NP-hard to approximate, it is also conjectured to have no FPRAS [\[DGGJ04\]](#). On the other hand, the same problem in undirected graphs is less clear. The only prior result is due to Karger [\[Kar01\]](#), where an FPRAS is presented if the minimum $s - t$ cut has the same order as the global minimum cut. We show that in undirected graphs, the problem is $\#\text{BIS}$ -hard as well, and thus is unlikely to admit an FPRAS.

Theorem 3. *For any $p \in (0, 1)$, if undirected $s - t$ unreliability at fixed open probability p for all edges admits an FPRAS, then $\#\text{BIS}$ admits an FPRAS.*

Indeed, Provan and Ball [\[PB83\]](#) already gave a parsimonious reduction from $\#\text{BIS}$ to the number of $s - t$ minimum cuts. The “ground states” for $s - t$ unreliability are exactly these min-cuts, and thus we just need a boosting argument so that these states contribute a significant portion of the unreliability. A full proof is given in [Section 5](#) for completeness. Note that if we consider $s - t$ *minimal* cuts, the problem becomes NP-hard [\[GGL16\]](#).

The $\#\text{BIS}$ -hardness of [Theorem 3](#) and in [\[FG24\]](#) leaves open the possibility of $s - t$ unreliability being either NP-hard or $\#\text{BIS}$ -easy. We find both possibilities to be plausible and would not make a conjecture here.

1.1 Proof overview

Given a graph $G = (V, E)$, it is natural to define a so-called Gibbs distribution, whose states are $F \subseteq E$ such that s can reach t in (V, F) , and the probability of F is proportional to the probability of F being open in the product distribution (bond percolation). Approximating $\text{Rel}(s, t)$ can reduce to sampling from this Gibbs distribution, and the standard way to do so is Glauber dynamics. Since this recipe works well for all-terminal reliability (see, e.g. [\[CGZZ24\]](#)), it is tempting to try it for two-terminal reliability as well. However, it fails quite badly for a simple example where s and t are connected by two long paths. If either path is open, s can reach t , but it is exponentially unlikely to switch from one path to another. It implies straightforward Glauber dynamics mixes slowly on graphs like this.

To avoid this rigidity, one idea is to enlarge the state space to include both a subset $F \subseteq E$ of arcs and a marked vertex u , so that s can always reach u in (V, F) . If we assign a uniform weight on all such states, then the resulting distribution biases towards the vertices u whose $\text{Rel}(s, u)$ is large. To ensure that we can sample (t, F) with an at least inverse polynomial probability, we will need to add a correction term of roughly, say, $1/\text{Rel}(s, u)$ for each possible marked vertex u . However, our original goal is to estimate $\text{Rel}(s, t)$, so this sounds like we are going in circles.

The key idea to salvage the above is to use simulated annealing. Instead of open probability p_e for each e , we first consider $p_e = 1$ for all e . In this case it is trivial to evaluate $\text{Rel}(s, u)$ for all u . Then, we gradually reduce each p_e by a small multiplicative factor, which is essentially the step of adjusting temperatures in simulated annealing. This ensures that $\text{Rel}(s, u)$'s between two adjacent temperatures change only by a constant factor. Then, we can use the estimates of $\text{Rel}(s, u)$ of the previous temperature to define the distribution of the next temperature, which has only a constant multiplicative error to what we desire.

With the target distribution in hand, the remaining question is how to sample from it. We consider a Markov chain with two types of transitions, either to update the arc set or to update the marked vertex together with an incident arc. We use the multicommodity flow method [Sin92] to bound the congestion of this chain. The key is to route the probability mass between any pair of states while maintaining a low congestion. In fact, the pioneering work of canonical paths by Jerrum and Sinclair [JS89] has already shown us how to distribute the probability mass evenly when the global changes are flipping paths and cycles. In our case, for any two states (u, F) and (u', F') , there is a canonical way of choosing a (backwards) path from u to s and then from s to u' . We use the idea of [JS89] to distribute the flow to flip arcs along this path. However, one complication is that when some arc is removed, the marked vertex needs to move too. This is allowed in our enlarged state space, and indeed appears to be the main benefit of considering it. Our (random) path from (u, F) to (u', F') always goes through an intermediate state (s, U) , where U is a random subgraph from the product distribution. This random U distributes the weights somewhat evenly in the state space. Intuitively, there is no constraint on the arcs once the marked vertex is s , so that arcs can be freely changed. Some extra care is needed here and our constructed flow actually interleaves between the movement of the marked vertex and other arc flips. This is the most involved step of the algorithm, and the details can be found in Section 3.4.

Overall, the high level idea of using simulated annealing together with Markov chain Monte Carlo is very similar to the non-negative permanent approximation algorithm by Jerrum, Sinclair, and Vigoda [JSV04]. However, the details are quite different and the construction of the flow paths appears to be new. We hope this framework to find more applications in the future.

2 Proof of Theorem 1

In this section, we first give a high-level outline of the algorithm in Theorem 1. The main algorithm and the annealing is explained in Section 2.1 and the Markov chain step in Section 2.2. The analysis of the Markov chain is in Section 3.

Let $G = (V, E)$ be a directed graph with two distinguished vertices s, t . If $s = t$, then $\text{Rel}(s, t) = 1$ and there is nothing to estimate, so assume that $s \neq t$. If G contains self-loops, they can be safely removed. If G contains parallel arcs, say with open probability p_1 and p_2 , then they can be merged into a single arc with an effective open probability $p_1 + p_2 - p_1p_2$. Thus we may assume that G is simple. This reduction never decreases the value of $p_{\min}$. It does not affect the running time upper bound in Theorem 1. Moreover, we assume throughout that s can reach every vertex of G . This is without loss of generality: no path starting at s can use an arc incident to a vertex that s cannot reach, so every such vertex and arc can be deleted without changing $\text{Rel}(s, t)$, and if t is deleted then $\text{Rel}(s, t) = 0$. Let $p_e \in (0, 1]$ be the open probability of each arc $e \in E$. Let $\mathbf{p} = (p_e)_{e \in E}$ denote the vector of open probabilities. Let $\mathcal{D}_{\mathbf{p}}$ be the product measure with Bernoulli distribution with

parameter $\mathbf{p}$ (namely bond percolation):

$$\forall F \subseteq E, \quad \mathcal{D}_{\mathbf{p}}(F) := \prod_{e \in F} p_e \prod_{e \in E \setminus F} (1 - p_e).$$

For any vertex $v \in V$, define the collection of arc sets

$$\Omega_v := \{F \subseteq E : s \text{ can reach } v \text{ via arcs in } F\}.$$

Clearly, the probability that s can reach v , denoted by $q_v(\mathbf{p})$, is given by

$$q_v(\mathbf{p}) := \sum_{F \in \Omega_v} \mathcal{D}_{\mathbf{p}}(F).$$

Note that q_v is a function of $\mathbf{p}$, and we write q_v if $\mathbf{p}$ is clear from the context. In particular, $q_t = \text{Rel}(s, t)$ is the value of interest and $q_s = 1$. As s can reach every vertex, $q_v > 0$ for every $v \in V$.

It is convenient to consider the more general problem of approximating q_v for all vertices $v \in V$ simultaneously. Define the following set of pairs

$$\Omega = \{(v, F) \mid v \in V, F \in \Omega_v\}.$$

We regard every pair $X = (v, F) \in \Omega$ as a state, and call v the *marked vertex* of X and F its arc set. Suppose one can draw a random state $X = (v, F) \in \Omega$ with probability proportional to $\mathcal{D}_{\mathbf{p}}(F)$. The states whose marked vertex equals v have total weight $\sum_{F \in \Omega_v} \mathcal{D}_{\mathbf{p}}(F) = q_v$, so the marked vertex of X equals v with probability exactly $q_v / \sum_{u \in V} q_u$. Given N independent samples of X , let N_v be the number whose marked vertex is v . Then $\mathbb{E}[N_v/N] = q_v / \sum_{u \in V} q_u$ for every $v \in V$, and since $q_s = 1$, the ratio N_v/N_s estimates q_v with the normalising constant $\sum_{u \in V} q_u$ cancelled. However, this estimator is useless as it stands. If q_v is exponentially small, then so is the probability that a single sample has marked vertex v , and one needs exponentially many samples before N_v is nonzero. The solution is to reweight the vertices so that none of them is rare. We introduce the following distribution.

Definition 4. Given a vector $\mathbf{p}$ of open probabilities and a vector of positive vertex weights $\mathbf{c} = (c_v)_{v \in V}$, define a distribution $\pi_{\mathbf{c}, \mathbf{p}}$ over Ω by

$$\forall (v, F) \in \Omega, \quad \pi_{\mathbf{c}, \mathbf{p}}(v, F) = \frac{c_v \mathcal{D}_{\mathbf{p}}(F)}{\sum_{u \in V} c_u q_u(\mathbf{p})} \propto c_v \mathcal{D}_{\mathbf{p}}(F).$$

Suppose we can draw $X = (v, F) \sim \pi_{\mathbf{c}, \mathbf{p}}$ and let N_v be the number of samples whose marked vertex is v among N independent samples of X . Then $\mathbb{E}[N_v/N] = c_v q_v / \sum_{u \in V} c_u q_u$, and since $q_s = 1$, the normalising constant once again cancels in the ratio: one can estimate q_v by $\frac{c_s}{c_v} \cdot \frac{N_v}{N_s}$. The point of the weights is that they can flatten the distribution of the marked vertex. Ideally, one would set $c_v = 1/q_v$ for all $v \in V$, in which case the marked vertex is uniform over V , as it equals v with probability $c_v q_v / \sum_{u \in V} c_u q_u = 1/n$ for every v , where $n = |V|$. No value of the marked vertex is then rare, and N can be taken polynomially large.

The ideal weights are of course the very quantities to be approximated, and cannot be set in advance. It turns out that knowing $1/q_v$ up to a constant factor is enough. The following lemma shows that such weights already allow all the q_v to be estimated within relative error ε , and that sampling from $\pi_{\mathbf{c}, \mathbf{p}}$ can be done efficiently by a Markov chain.

Lemma 5. Suppose we are given open probabilities $\mathbf{p} = (p_e)_{e \in E}$ and positive vertex weights $\mathbf{c} = (c_v)_{v \in V}$ such that $\frac{1}{4q_v(\mathbf{p})} \leq c_v \leq \frac{4}{q_v(\mathbf{p})}$ for all $v \in V$.

1. There is a Markov chain on Ω with unique stationary distribution $\pi_{\mathbf{c}, \mathbf{p}}$. Independently draw $F_0 \sim \mathcal{D}_{\mathbf{p}}$ and start the chain from the random state $X_0 = (s, F_0)$. For any $\eta \in (0, 1)$, after running

$$T = O\left(nm^2 \log\left(\frac{n}{\eta}\right)\right)$$

steps, it outputs a random $X \in \Omega$ satisfying $d_{\text{TV}}(X, \pi_{\mathbf{c}, \mathbf{p}}) \leq \eta$, where d_{TV} is the total variation distance. Moreover, for every $\zeta \in (0, 1)$, the simulation of T -step Markov chain takes time

$$O\left(n^2m^2 \log\left(\frac{n}{\eta}\right) + m \log\left(\frac{1}{\zeta}\right)\right)$$

with probability at least $1 - \zeta$.

2. For any $\varepsilon, \delta \in (0, 1)$, using $N = O\left(\frac{n}{\varepsilon^2} \log\left(\frac{n}{\delta}\right)\right)$ samples of the above Markov chain, each with $\eta = \frac{\delta}{4N}$, one can produce random $(\hat{q}_v)_{v \in V}$ such that

$$\Pr[\forall v \in V, (1 - \varepsilon)q_v(\mathbf{p}) \leq \hat{q}_v \leq (1 + \varepsilon)q_v(\mathbf{p})] \geq 1 - \delta$$

in worst-case time

$$O\left(\frac{n^3m^2}{\varepsilon^2} \log\left(\frac{n}{\varepsilon\delta}\right) \log\left(\frac{n}{\delta}\right)\right).$$

The Markov chain is described in [Section 2.2](#). Once Item (1) is established, Item (2) follows from the estimator discussed earlier. The proof of Item (1) is deferred to [Section 3](#), and the proof of Item (2) is deferred to [Section 4](#). In [Section 2.1](#) next, we show how to use [Lemma 5](#) to approximate the two-terminal reliability.

2.1 The main algorithm

Recall that $\mathbf{p} = (p_e)_{e \in E}$ is the vector of open probabilities. To apply [Lemma 5](#), we need to construct a vector $\mathbf{c} = (c_v)_{v \in V}$ satisfying $\frac{1}{4q_v(\mathbf{p})} \leq c_v \leq \frac{4}{q_v(\mathbf{p})}$ for all $v \in V$. We call such a vector $\mathbf{c}$ admissible. This can be done by an annealing procedure on open probabilities. The same idea was used by Jerrum, Sinclair, and Vigoda [[JSV04](#)] to approximate the permanent of a non-negative matrix.

We first record that the comparison between two nearby probability vectors depends only on the size of a reachability certificate, rather than on the total number of open arcs.

Lemma 6. Let $\mathbf{p}' = (p'_e)_{e \in E}$ and $\mathbf{p} = (p_e)_{e \in E}$ be two vectors of open probabilities. If, for some $\beta \in (0, 1]$, it holds that $\beta p_e \leq p'_e \leq p_e$ for all $e \in E$, then, for every $v \in V$, $q_v(\mathbf{p}') \geq \beta^{n-1} q_v(\mathbf{p})$.

Proof. Sample $F \sim \mathcal{D}_{\mathbf{p}}$. Conditional on F , independently retain every arc $e \in F$ with probability p'_e/p_e , and let F' be the resulting arc set. The indicators of membership in F' are independent, and each arc e belongs to F' with probability $p_e(p'_e/p_e) = p'_e$. Hence $F' \sim \mathcal{D}_{\mathbf{p}'}$.

For every $F \in \Omega_v$, fix a simple directed path $P(F, v)$ from s to v contained in F . Such a path contains at most $n - 1$ arcs. Given $F \in \Omega_v$, the probability that all arcs of $P(F, v)$ survive the thinning is at least

$$\prod_{e \in P(F, v)} \frac{p'_e}{p_e} \geq \beta^{|P(F, v)|} \geq \beta^{n-1}.$$

Whenever this path survives, $F' \in \Omega_v$. Therefore $q_v(\mathbf{p}') = \Pr[F' \in \Omega_v] \geq q_v(\mathbf{p})\beta^{n-1}$. $\square$

In our algorithm, we set the parameter

$$\beta := 1 - \frac{1}{4n}.$$

Definition 7. Define a sequence of open probabilities $\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \dots, \mathbf{p}^{(L)}$, where each $\mathbf{p}^{(i)} = (p_e^{(i)})_{e \in E}$ is a vector of open probabilities for the arcs $e \in E$. The sequence satisfies

- $\mathbf{p}^{(0)} = \mathbf{1}$ is the all-one vector and $\mathbf{p}^{(L)} = \mathbf{p}$ is the input open probabilities.
- for any $0 < i \leq L$ and any $e \in E$, $\beta p_e^{(i-1)} \leq p_e^{(i)} \leq p_e^{(i-1)}$.

The following lemma shows that one can construct a sequence of open probabilities that satisfies the above definition with a length polynomial in the input size. The proof is given in [Section 4](#).

Lemma 8. Let $p_{\min} = \min_{e \in E} p_e$. One can construct a sequence of open probabilities that satisfies [Definition 7](#) with a length $L = O\left(n \log\left(\frac{2}{p_{\min}}\right)\right)$ in time $O(mL)$.

Fix the underlying graph $G = (V, E)$, two terminals $s, t \in V$, and the sequence of open probabilities as in [Definition 7](#), and think of the open probabilities as gradually changing from $\mathbf{p}^{(0)} = \mathbf{1}$ to $\mathbf{p}^{(L)} = \mathbf{p}$. At the i -th stage, we consider the two-terminal reliability problem with the open probabilities $\mathbf{p}^{(i)}$, and abbreviate by $q_v^{(i)} = q_v(\mathbf{p}^{(i)})$ the probability that s can reach v at that stage. The point of [Definition 7](#) is that consecutive stages have comparable reachability probabilities, so that the reciprocals of the estimates of $q_v^{(i)}$ obtained at the i -th stage are admissible vertex weights c_v for the $(i + 1)$ -th stage. Formally, we have the following lemma.

Lemma 9. For every vertex $v \in V$ and every $1 \leq i \leq L$,

$$\frac{3}{4} \leq \frac{q_v^{(i)}}{q_v^{(i-1)}} \leq 1.$$

Proof. As $p_e^{(i)} \leq p_e^{(i-1)}$ for every $e \in E$, it is easy to verify $q_v^{(i)} \leq q_v^{(i-1)}$. For the reverse inequality, the definition of the schedule and [Lemma 6](#) give

$$q_v^{(i)} \geq \beta^{n-1} q_v^{(i-1)} \geq \left(1 - \frac{n-1}{4n}\right) q_v^{(i-1)} > \frac{3}{4} q_v^{(i-1)},$$

where the second inequality is Bernoulli's inequality. This completes the proof. $\square$

Proof of the main result Now we are ready to describe the algorithm of [Theorem 1](#). It computes estimates $\hat{q}_v^{(i)}$ of $q_v^{(i)}$ for every $v \in V$, one stage at a time for $i = 0, 1, \dots, L$, and outputs $\hat{q}_t^{(L)}$. At the 0-th stage, $p_e^{(0)} = 1$ for every $e \in E$, so $q_v^{(0)} = 1$ for every $v \in V$ and the algorithm takes $\hat{q}_v^{(0)} = 1$. Consider the i -th stage for $1 \leq i \leq L$, and assume inductively that the estimates of the previous stage satisfy

$$\frac{9}{10}q_v^{(i-1)} \leq \hat{q}_v^{(i-1)} \leq \frac{11}{10}q_v^{(i-1)} \quad (v \in V), \quad (1)$$

which holds trivially when $i = 1$. Set the vertex weights

$$c_v^{(i)} := 1/\hat{q}_v^{(i-1)} \quad (v \in V),$$

and call the algorithm in Item (2) of [Lemma 5](#) on the open probabilities $\mathbf{p}^{(i)}$ with the weights $\mathbf{c}^{(i)}$, using fresh randomness. These weights are admissible: combining (1) with [Lemma 9](#),

$$c_v^{(i)} q_v^{(i)} = \frac{q_v^{(i)}}{\hat{q}_v^{(i-1)}} \geq \frac{10}{11} \cdot \frac{q_v^{(i)}}{q_v^{(i-1)}} \geq \frac{10}{11} \cdot \frac{3}{4} > \frac{1}{4} \quad \text{and} \quad c_v^{(i)} q_v^{(i)} \leq \frac{10}{9} \cdot \frac{q_v^{(i)}}{q_v^{(i-1)}} \leq \frac{10}{9} < 4,$$

so that $\frac{1}{4q_v^{(i)}} \leq c_v^{(i)} \leq \frac{4}{q_v^{(i)}}$ for every $v \in V$, as [Lemma 5](#) requires. For $i < L$, the algorithm is called with $\varepsilon = \frac{1}{10}$ and $\delta = \frac{1}{10L}$, and its output satisfies (1) at the next stage with probability at least $1 - \frac{1}{10L}$. For $i = L$, it is called with the target relative error ε and with $\delta = \frac{1}{10}$. Since $\mathbf{p}^{(L)} = \mathbf{p}$, the returned value $\hat{q}_t^{(L)}$ estimates $q_t^{(L)} = \text{Rel}(s, t)$.

Let A_i be the event that the i -th stage meets its requirement, that is, (1) holds with i in place of $i-1$ if $i < L$, and $(1-\varepsilon)q_t^{(L)} \leq \hat{q}_t^{(L)} \leq (1+\varepsilon)q_t^{(L)}$ if $i = L$. As every stage uses fresh randomness, the two bounds above apply conditionally on the previous stages, and

$$\Pr \left[\bigcap_{i=1}^L A_i \right] = \prod_{i=1}^L \Pr \left[A_i \mid \bigcap_{j<i} A_j \right] \geq \left(1 - \frac{1}{10L} \right)^{L-1} \left(1 - \frac{1}{10} \right) \geq \frac{3}{4}.$$

The output is therefore within a factor $1 \pm \varepsilon$ of $\text{Rel}(s, t)$ with probability at least $\frac{3}{4}$.

It remains to bound the running time. The bound of [Lemma 5](#) applies uniformly at every stage. Each of the first $L-1$ stages costs

$$O\left(n^3 m^2 \log^2(nL)\right) = \tilde{O}\left(n^3 m^2\right),$$

and the last stage costs $\tilde{O}\left(\frac{n^3 m^2}{\varepsilon^2}\right)$. Since $L = O\left(n \log\left(\frac{2}{p_{\min}}\right)\right)$ by [Lemma 8](#), the total running time is

$$\tilde{O}\left(n^4 m^2 \log\left(\frac{2}{p_{\min}}\right) + \frac{n^3 m^2}{\varepsilon^2}\right),$$

which is at most the bound stated in [Theorem 1](#).

2.2 The Markov chain

We now describe the Markov chain used in [Lemma 5](#). To simplify the notation, we fix open probabilities $\mathbf{p} = (p_e)_{e \in E}$ and vertex weights $\mathbf{c} = (c_v)_{v \in V}$ such that $\frac{1}{4q_v} \leq c_v \leq \frac{4}{q_v}$ for all $v \in V$, where $q_v = q_v(\mathbf{p})$. We use π to denote the distribution $\pi_{\mathbf{c}, \mathbf{p}}$ over Ω . From now on we also restrict Ω to the support of π , discarding every pair (v, F) with $\mathcal{D}_{\mathbf{p}}(F) = 0$, so that $\pi(X) > 0$ for every $X \in \Omega$. This restriction is vacuous unless $p_e = 1$ for some arc e , and it is harmless: the discarded states carry no probability mass. Moreover, if $F_0 \sim \mathcal{D}_{\mathbf{p}}$, then $\mathcal{D}_{\mathbf{p}}(F_0) > 0$ almost surely, so the random state (s, F_0) used below belongs to the restricted state space almost surely.

The Markov chain for sampling from $\pi_{\mathbf{c}, \mathbf{p}}$ is described as follows. It starts from a random initial state X_0 with some distribution μ on Ω . For the sampling algorithm, we independently draw $F_0 \sim \mathcal{D}_{\mathbf{p}}$ and take $X_0 = (s, F_0)$. The chain uses two transition kernels K_1 and K_2 . At the k -th step, with half probability, stay at the current state $X_{k+1} = X_k$; with probability $\frac{1}{4}$, it uses K_1 to update X_k to X_{k+1} , and with probability $\frac{1}{4}$, it uses K_2 . In other words, the overall transition kernel K is given by

$$K = \frac{1}{2}I + \frac{1}{4}K_1 + \frac{1}{4}K_2, \quad (2)$$

where I is the identity. Let $X_k = (v, F)$ be the current state. The kernel K_1 updates the marked vertex v , but it may also change the arc set F . The other kernel K_2 updates the arc set F . The detailed definitions are as follows.

Metropolis update K_1 : pick an arc $e = (a, b) \in E$ uniformly at random, and propose a candidate state X'_{k+1} as follows.

- If $v = a$, then with half probability propose $X'_{k+1} = (b, F \cup \{e\})$ and with half probability propose $X'_{k+1} = X_k$.
- If $v = b$ and $e \in F$, then with half probability propose $X'_{k+1} = (a, F \cup \{e\})$ and with half probability propose $X'_{k+1} = (a, F \setminus \{e\})$.
- In all other cases ($v \notin \{a, b\}$ or ($v = b$ and $e \notin F$)), propose $X'_{k+1} = X_k$.

With probability $\min \left\{ 1, \frac{\pi(X'_{k+1})}{\pi(X_k)} \right\}$, accept the proposal and set $X_{k+1} = X'_{k+1}$; otherwise, reject the proposal and set $X_{k+1} = X_k$. We remark that the proposal X'_{k+1} can be infeasible ($X'_{k+1} \notin \Omega$), in which case $\pi(X'_{k+1}) = 0$ and the proposal is always rejected.

Heat-bath update K_2 : pick an arc $e \in E$ uniformly at random, and let $F_1 = F \cup \{e\}$ and $F_0 = F \setminus \{e\}$. If s can reach v via arcs in F_0 , then let $X_{k+1} = (v, F_1)$ with probability p_e and $X_{k+1} = (v, F_0)$ with probability $1 - p_e$. Otherwise, let $X_{k+1} = (v, F_1)$.

For any Markov chain P over some finite state space Ω , it is called *irreducible* if for any two states $X, Y \in \Omega$, there exists a positive integer t such that $K^t(X, Y) > 0$. Moreover, it is *aperiodic* if for any state $X \in \Omega$, the greatest common divisor of the set of all positive integers t such that $K^t(X, X) > 0$ is 1, and it is *reversible* with respect to π if for any two states $X, Y \in \Omega$, the detailed balance condition $\pi(X)K(X, Y) = \pi(Y)K(Y, X)$ holds. If a chain is irreducible, aperiodic, and reversible, then it converges to a unique stationary distribution. See [\[LP17\]](#).

Proposition 10. *The Markov chain K defined in (2) is irreducible, aperiodic, and reversible with respect to π .*

Proposition 10 is verified in Section 3.2. Thus the chain K has π as its unique stationary distribution, and the distribution of X_k converges to π as $k \rightarrow \infty$. Given a random initial state X_0 with distribution μ on Ω and an accuracy $\eta \in (0, 1)$, define the mixing time of the Markov chain K as the smallest number of steps after which the chain is within total variation distance η of π , that is,

$$T_{\text{mix}}(\mu, \eta) = \min \left\{ k \geq 0 : d_{\text{TV}}(\mu K^k, \pi) \leq \eta \right\},$$

where μK^k is the distribution of X_k . We call μ an M -warm start with respect to π if

$$\forall Y \in \Omega, \quad \mu(Y) \leq M\pi(Y).$$

The analysis in Section 3 shows the following bound on the mixing time.

Theorem 11. *Let $G = (V, E)$ be a directed graph with $n = |V|$ vertices and $m = |E|$ arcs such that the source s can reach every vertex $v \in V$. Let $\mathbf{p} = (p_e)_{e \in E}$ be open probabilities with $p_e \in (0, 1]$ for all $e \in E$, and let $\mathbf{c} = (c_v)_{v \in V}$ be vertex weights such that $\frac{1}{4q_v} \leq c_v \leq \frac{4}{q_v}$ for all $v \in V$. For every initial distribution μ that is M -warm with respect to π and every $\eta \in (0, 1)$, the mixing time of the Markov chain K satisfies*

$$T_{\text{mix}}(\mu, \eta) = O \left(nm^2 \left(\log M + \log \left(\frac{1}{\eta} \right) \right) \right).$$

Theorem 11 is proved in Section 3 by analyzing the spectral gap of the transition matrix.

Now, assume the mixing time is known. We are ready to prove the first item in Lemma 5.

Proof of Item (1) of Lemma 5. The Markov chain is K as defined in (2). Independently draw $F_0 \sim \mathcal{D}_{\mathbf{p}}$ and take $X_0 = (s, F_0)$. This is a legal state almost surely because s can always reach itself and $\mathcal{D}_{\mathbf{p}}(F_0) > 0$ almost surely. Let μ denote the distribution of X_0 , and write $Z = \sum_{u \in V} c_u q_u$. For every $(v, F) \in \Omega$, we have

$$\mu(v, F) = \begin{cases} \mathcal{D}_{\mathbf{p}}(F), & v = s, \\ 0, & v \neq s. \end{cases}$$

Recall that $q_s = 1$, so $c_s \geq \frac{1}{4}$, while $c_u q_u \leq 4$ for every $u \in V$ gives $Z \leq 4n$. Consequently,

$$\forall (s, F) \in \Omega, \quad \frac{\mu(s, F)}{\pi(s, F)} = \frac{Z}{c_s} \leq 16n,$$

and the same pointwise domination is trivial for states whose marked vertex is not s . Thus μ is a $16n$ -warm start with respect to π . By Theorem 11, running the chain for

$$T_{\text{mix}}(\mu, \eta) = O \left(nm^2 \left(\log(16n) + \log \left(\frac{1}{\eta} \right) \right) \right) = O \left(nm^2 \log \left(\frac{n}{\eta} \right) \right)$$

steps outputs a random $X \in \Omega$ with $d_{\text{TV}}(X, \pi) \leq \eta$.

Implementation and running time. A trivial implementation of each transition takes $O(m)$ time: as every vertex is reachable from s , we have $m \geq n - 1$, and one breadth-first search tests the required reachability condition in $O(m)$ time. We next give an implementation with a sharper bound for a sequence of transitions.

Store the static outgoing adjacency lists of G . Let (v, F) denote the current state. Represent F by a bit vector in $\{0, 1\}^E$, and maintain a simple directed path P from s to the current marked vertex v using arcs of F , together with arrays recording the vertices and arcs on P . Initially, when $X_0 = (s, F_0)$, take P to be the length-zero path at s . Consider first a K_2 update of an arc e . If $e \notin P$, then P remains a witness that s reaches v in $F \setminus \{e\}$, so no search is needed. If $e \in P$, run breadth-first search in $F \setminus \{e\}$. If it finds another path from s to v , use that path as the new witness; otherwise K_2 must retain e , and the old path remains valid.

For a K_1 update, suppose first that the selected arc is $e = (v, w)$. The candidate state with marked vertex w is automatically feasible: if the proposal is accepted, append e to P when w is not on P , and use the prefix ending at w when w is on P . Now suppose that the selected arc is $e = (u, v) \in F$. Run a breadth-first search in the arc set of the proposed state to test whether s can reach u . If the proposal is feasible and accepted, use the path returned by the search as the new witness; otherwise, retain P . Thus, a K_1 update only requires a search in this last case.

Apart from a breadth-first search, the path can be updated in $O(n)$ time. Since F and F' differ on at most one arc, the Metropolis acceptance probability can be computed in $O(1)$ time.

In the above implementation, only some transitions require a breadth-first search. Recall that T is the total number of transitions we simulate. For $k \in [T]$, let $I_k \in \{0, 1\}$ indicate that the k -th transition performs a breadth-first search, and let $\mathcal{H}_{k-1}$ contain the full history before that transition. Conditional on this history, the current witness path has at most $n - 1$ arcs. A K_2 step performs a search only if its uniformly selected arc lies on this path. A K_1 step performs a search only if its selected arc is an incoming arc of the current marked vertex. Since G is simple, the latter also has at most $n - 1$ possibilities. Recalling the coefficients of K_1 and K_2 in (2), we obtain

$$\mathbb{E}[I_k \mid \mathcal{H}_{k-1}] \leq \frac{n-1}{4m} + \frac{n-1}{4m} \leq \frac{n-1}{2m}. \quad (3)$$

The random variables $I_1, \dots, I_T$ need not be independent, but (3) gives

$$\mathbb{E}[\exp(I_k) \mid \mathcal{H}_{k-1}] \leq 1 + (e-1) \frac{n-1}{2m} \leq \exp\left((e-1) \frac{n-1}{2m}\right).$$

Iterating this conditional bound yields

$$\mathbb{E}\left[\exp\left(\sum_{k=1}^T I_k\right)\right] \leq \exp\left((e-1) \frac{(n-1)T}{2m}\right).$$

Consequently, Markov's inequality implies that, for every $\zeta \in (0, 1)$,

$$\Pr\left[\sum_{k=1}^T I_k > \frac{e(n-1)T}{2m} + \log\left(\frac{1}{\zeta}\right)\right] \leq \zeta. \quad (4)$$

Let $\text{Time}(T)$ denote the total running time of one simulation consisting of initialisation followed by T transitions. Drawing F_0 and initialising the data structure take $O(m)$ time. Since a transition

costs $O(n)$ time apart from a possible $O(m)$ breadth-first search, (4) shows that, for a sufficiently large universal constant C ,

$$\Pr \left[\text{Time}(T) > C \left(m + nT + m \log \left(\frac{1}{\zeta} \right) \right) \right] \leq \zeta. \quad (5)$$

Finally, substituting $T = O(nm^2 \log(n/\eta))$ into (5) proves the running-time bound in Item (1). $\square$

3 Analysis of the Markov chain

In this section, we prove [Proposition 10](#) and [Theorem 11](#). This section is organized as follows. In [Section 3.1](#), we review some preliminaries of the Markov chain analysis. In [Section 3.2](#), we verify [Proposition 10](#), namely the irreducibility, aperiodicity, and reversibility of the chain K . In [Section 3.3](#), we prove a lower bound on the spectral gap of K , which implies the mixing time bound in [Theorem 11](#).

Intuition of rapid mixing Before turning to the analysis, we give some intuition why K mixes rapidly. Consider the most natural chain for the problem. The quantity to be estimated is $q_t = \mathcal{D}_p(\Omega_t)$, and the direct approach is to sample an arc set F from $\mathcal{D}_p$ conditioned on the event $F \in \Omega_t$, namely conditioned on s reaching t via the open arcs. The standard chain is the *heat-bath Glauber dynamics* on the state space Ω_t : from a current arc set $F \in \Omega_t$, pick an arc $e \in E$ uniformly at random and resample its membership according to the conditional marginal given the other arcs, keeping $F \setminus \{e\}$ fixed. This is precisely the kernel K_2 with the marked vertex fixed at t .

This natural chain can mix exponentially slowly. Suppose that G consists of two vertex-disjoint s - t paths of length ℓ , and that every arc is open with probability $\frac{1}{2}$. Reaching t requires at least one of the two paths to be entirely open, so $\Omega_t = A \cup C$, where A and C are the sets of arc sets in which the upper, respectively the lower, path is entirely open; see [Figure 2](#). As the dynamics changes one arc at a time, it cannot leave A before the lower path has become entirely open, so every trajectory from A to C has to pass through $B = A \cap C$. Such a bottleneck forces the mixing time to be $2^{\Omega(\ell)}$, exponential in the size of the graph.

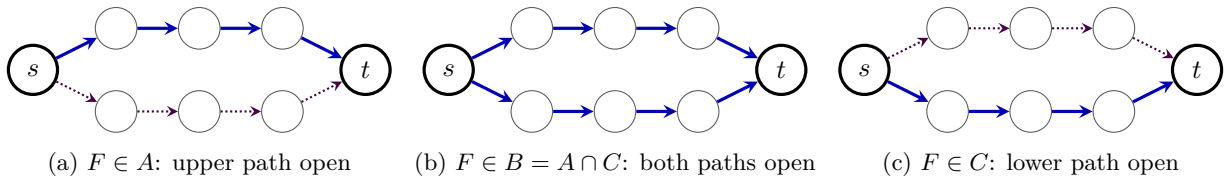


Figure 2: The Glauber dynamics can pass from A to C only through the rare set B . Solid arcs are open, and dotted arcs are subject to no constraint.

Our chain circumvents this issue by letting the marked vertex move. Recall that a state is a pair (v, F) in which the marked vertex v is reachable from s via the arc set F , and that the kernel K_1 updates the marked vertex. Consider again the state (t, F) in which the upper path is open, and let w be the predecessor of t on that path. Then K_1 can move the marked vertex from t back to w , updating the arc e between them along the way, as illustrated in [Figure 3](#). The reachability constraint travels with the marked vertex: in the new state only w has to be reachable from s .

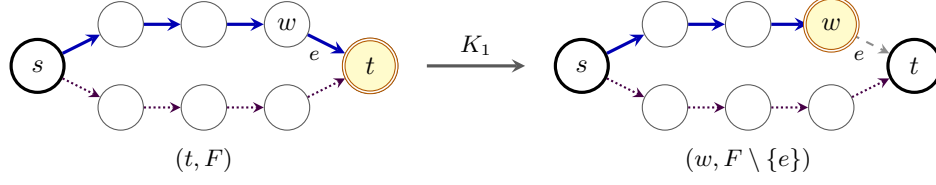


Figure 3: One step of K_1 , retreating the marked vertex from t to w .

Consider now the ideal choice of vertex weights $c_v = 1/q_v$. Then the marked vertex is uniform under π : for every $v \in V$, the states whose marked vertex is v have total stationary mass $1/n$. The K_1 moves allow the reachability requirement to move locally along the underlying undirected graph, while K_2 refreshes arcs that are not forced by the current requirement. In the two-path example, the marked vertex can retreat along the upper path from t to s , so that the old witness path is progressively no longer required. When the marked vertex is s , the reachability constraint is vacuous, the conditional distribution of the arc set is $\mathcal{D}_p$, and K_2 is the ordinary product heat-bath dynamics. The chain can therefore forget the upper witness path and subsequently grow towards t along the lower path, without requiring both paths to be open simultaneously. Thus the relevant gateway is the entire collection of states whose marked vertex is s , which has stationary mass $1/n$, rather than the exponentially rare set B . The same layer also provides a warm start for the sampling algorithm: draw $F \sim \mathcal{D}_p$ and mark s . This is exactly π conditioned on the marked vertex being s , and under admissible vertex weights it is a $16n$ -warm start with respect to π . In the following analysis, the multicommodity-flow argument in [Section 3.3](#) and [Section 3.4](#) makes this intuition rigorous.

We also note that in the two-path example above, if we start from s , it is roughly an unbiased simple random walk to reach t , which would take time $\Omega(\ell^2)$. The update rate is roughly $1/\ell$, making the overall mixing time $\Omega(\ell^3)$ to drop the total variation distance error down to $O(1/n)$. Thus there is a $\Omega(n^3)$ lower bound on the mixing time of K , which is a polynomial factor away from the upper bound in [Theorem 11](#). It is easy to see that this mixing time can be improved to $\Theta(n^2)$ if we change K_1 to update only an adjacent arc instead of a uniformly at random arc. However, even for that version there is still an $\Omega(n^3)$ mixing example, and there is no improvement on our mixing time upper bound as the bottleneck from K_2 is unchanged. For some further discussion on the lower bound side of the mixing time, see [Section 3.5](#).

3.1 Preliminaries of Markov chain analysis

Let π be a distribution over a finite set Ω with $\pi(x) > 0$ for every $x \in \Omega$, and let K be a transition kernel on Ω that is irreducible and reversible with respect to π . Assume in addition that K is *lazy*, namely that in every step it stays at the current state with probability at least $\frac{1}{2}$; in particular K is aperiodic. Reversibility makes K self-adjoint with respect to the inner product $\langle f, g \rangle_\pi = \sum_{x \in \Omega} \pi(x) f(x) g(x)$ on $\mathbb{R}^\Omega$, so all eigenvalues of K are real; laziness makes them nonnegative. The $|\Omega|$ eigenvalues of K can thus be arranged as $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_{|\Omega|} \geq 0$ and the *spectral gap* of K is defined as $\text{gap}(K) = 1 - \lambda_2 \in (0, 1]$.

Lemma 12 (Spectral-gap mixing bound; cf. [\[LP17, Theorem 12.4\]](#)). *Let K be as above. For every*

probability distribution μ on Ω and every integer $t \geq 0$,

$$d_{\text{TV}}(\mu K^t, \pi) \leq \frac{1}{2}(1 - \text{gap}(K))^t \left(\sum_{x \in \Omega} \frac{\mu(x)^2}{\pi(x)} - 1 \right)^{1/2}. \quad (6)$$

Consequently, if μ is an M -warm start with respect to π , then

$$d_{\text{TV}}(\mu K^t, \pi) \leq \frac{1}{2} \sqrt{M-1} e^{-t \text{gap}(K)}. \quad (7)$$

Proof. Let $h = \mu/\pi - 1$, viewed as a function in $\mathbb{R}^\Omega$. Then $\langle h, 1 \rangle_\pi = 0$. Reversibility gives

$$\frac{(\mu K^t)(y)}{\pi(y)} - 1 = (K^t h)(y) \quad (y \in \Omega).$$

Since K is lazy, all its eigenvalues are nonnegative, and its operator norm on the subspace orthogonal to the constants is $\lambda_2 = 1 - \text{gap}(K)$. Hence

$$2d_{\text{TV}}(\mu K^t, \pi) = \sum_{y \in \Omega} \pi(y) \left| (K^t h)(y) \right| \leq \|K^t h\|_{2,\pi} \leq (1 - \text{gap}(K))^t \|h\|_{2,\pi}.$$

Moreover,

$$\|h\|_{2,\pi}^2 = \sum_{x \in \Omega} \pi(x) \left(\frac{\mu(x)}{\pi(x)} - 1 \right)^2 = \sum_{x \in \Omega} \frac{\mu(x)^2}{\pi(x)} - 1,$$

which proves (6). If μ is M -warm, then

$$\sum_{x \in \Omega} \frac{\mu(x)^2}{\pi(x)} \leq M \sum_{x \in \Omega} \mu(x) = M.$$

Together with $1 - \text{gap}(K) \leq e^{-\text{gap}(K)}$, this proves (7). $\square$

The spectral gap can be captured by the Poincaré inequality. For a kernel K reversible with respect to π , the *Dirichlet form* of a function $f: \Omega \rightarrow \mathbb{R}$ is

$$\mathcal{E}_K(f, f) = \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) K(x, y) (f(x) - f(y))^2,$$

and the variance of f with respect to π is

$$\text{Var}_\pi(f) = \sum_{x \in \Omega} \pi(x) (f(x) - \mathbb{E}_\pi[f])^2 = \frac{1}{2} \sum_{x, y \in \Omega} \pi(x) \pi(y) (f(x) - f(y))^2. \quad (8)$$

The spectral gap is exactly the largest constant γ for which the Poincaré inequality $\gamma \text{Var}_\pi(f) \leq \mathcal{E}_K(f, f)$ holds for every f , that is [LP17, Lemma 13.7 and Remark 13.8],

$$\text{gap}(K) = \inf_{f: \text{Var}_\pi(f) \neq 0} \frac{\mathcal{E}_K(f, f)}{\text{Var}_\pi(f)}. \quad (9)$$

Lower bounding the spectral gap therefore amounts to establishing a Poincaré inequality: if $\text{Var}_\pi(f) \leq C \cdot \mathcal{E}_K(f, f)$ for every $f: \Omega \rightarrow \mathbb{R}$, then $\text{gap}(K) \geq \frac{1}{C}$.

A general method for establishing such a spectral gap bound is the canonical path technique of Jerrum and Sinclair [JS89] and Diaconis and Stroock [DS91], in the multicommodity flow formulation of Sinclair [Sin92]. View the Markov chain as a directed state graph on the vertex set Ω , in which an ordered pair $\eta = (z, z')$ of distinct states forms a transition whenever $K(z, z') > 0$, and equip this transition with the *capacity*

$$Q(\eta) = Q(z, z') := \pi(z)K(z, z'),$$

which is symmetric in z and z' by reversibility. A *path* from x to y is a sequence of transitions $\gamma = (\eta_1, \dots, \eta_\ell)$ with $\eta_i = (z_{i-1}, z_i)$, $z_0 = x$, $z_\ell = y$, and $z_0, \dots, z_\ell$ pairwise distinct; its *length* $|\gamma| = \ell$ is the number of transitions it uses. Let $\mathcal{P}_{x,y}$ be the set of simple paths from x to y , and let $\mathcal{P} = \bigcup_{x \neq y} \mathcal{P}_{x,y}$.

Definition 13 (flow and congestion). A *flow* for K is a function $\Gamma: \mathcal{P} \rightarrow \mathbb{R}_{\geq 0}$ that routes $\pi(x)\pi(y)$ units of demand from x to y for every ordered pair of distinct states, namely

$$\forall x \neq y \in \Omega, \quad \sum_{\gamma \in \mathcal{P}_{x,y}} \Gamma(\gamma) = \pi(x)\pi(y).$$

The *congestion* of Γ is

$$\rho(\Gamma) = \max_{\eta} \frac{1}{Q(\eta)} \sum_{\gamma \in \mathcal{P}: \eta \in \gamma} \Gamma(\gamma) |\gamma|,$$

where the maximum ranges over all transitions η and $\eta \in \gamma$ means that γ traverses η .

It is well known that a flow of small congestion certifies a Poincaré inequality.

Lemma 14 ([Sin92]). *For every flow Γ and every $f: \Omega \rightarrow \mathbb{R}$, it holds that $\text{gap}(K) \geq \frac{1}{\rho(\Gamma)}$.*

3.2 Basic properties of the Markov chain

We now verify [Proposition 10](#). Two facts are used throughout. First, π is strictly positive on every state of Ω : indeed $\pi(v, F) \propto c_v \mathcal{D}_p(F)$, where $c_v \geq \frac{1}{4q_v} > 0$ by assumption and $\mathcal{D}_p(F) > 0$ because Ω has been restricted to the support of π in [Section 2.2](#). Second, G is simple, so the two endpoints of an arc are distinct and every proposal of K_1 that differs from the current state changes the marked vertex.

Proof of Proposition 10. Aperiodicity is immediate, since $K(X, X) \geq \frac{1}{2} > 0$ for every $X \in \Omega$, so that 1 belongs to the set of return times of every state. For irreducibility, it is straightforward to verify that any $(u, F) \in \Omega$ can reach the state (s, E) using transitions K_1 and K_2 and the state (s, E) can reach any $(u, F) \in \Omega$ by reversing the transitions.

We turn to reversibility. The identity kernel is reversible with respect to every distribution, so by the decomposition $K = \frac{1}{2}I + \frac{1}{4}K_1 + \frac{1}{4}K_2$ it suffices to prove that K_1 and K_2 are both reversible with respect to π .

Let R be the proposal kernel of K_1 , so that $K_1(X, Y) = R(X, Y) \min\{1, \frac{\pi(Y)}{\pi(X)}\}$ for every $Y \neq X$. We claim that R is symmetric. Fix two distinct states $X = (v, F)$ and $Y = (u, F')$. As noted above,

a proposal that differs from the current state moves the marked vertex v to the other endpoint of the chosen arc, so $R(X, Y) = 0$ unless $u \neq v$, which we assume from now on. By the two cases in the definition of K_1 , each of the relevant proposals is made with probability $\frac{1}{m} \cdot \frac{1}{2}$, and therefore

$$R(X, Y) = \frac{N(X, Y)}{2m},$$

where $N(X, Y) \in \{0, 1, 2\}$ is the number of arcs $e \in E$ such that

- (a) $e = (v, u)$ and $F' = F \cup \{e\}$; or
- (b) $e = (u, v)$, $e \in F$, and $F' \in \{F, F \setminus \{e\}\}$.

Here (a) is the first case in the definition of K_1 and (b) is the second one. Note that $N(X, Y)$ can be as large as 2: this happens exactly when $F' = F$ and both (v, u) and (u, v) are in F . Now an arc e satisfies (a) for the ordered pair (X, Y) if and only if it satisfies (b) for (Y, X) . Indeed, if $e = (v, u)$ and $F' = F \cup \{e\}$, then $e \in F'$ and $F \in \{F', F' \setminus \{e\}\}$, which is (b) for (Y, X) ; conversely, if $e = (v, u)$, $e \in F'$ and $F \in \{F', F' \setminus \{e\}\}$, then $F \cup \{e\} = F'$, which is (a) for (X, Y) . Applying this equivalence gives $N(X, Y) = N(Y, X)$, and the claim follows. Consequently, for all $X \neq Y$,

$$\pi(X)K_1(X, Y) = R(X, Y) \min\{\pi(X), \pi(Y)\} = \pi(Y)K_1(Y, X), \quad (10)$$

which is the detailed balance condition for K_1 . Note that infeasible proposals cause no harm: they have $\pi(Y) = 0$ and are rejected with probability one.

For K_2 , fix an arc $e \in E$ and a state $X = (v, F) \in \Omega$, and write $F_1 = F \cup \{e\}$ and $F_0 = F \setminus \{e\}$. The update keeps the marked vertex v fixed and changes at most the membership of e , so the only possible transition between two distinct states is between (v, F_0) and (v, F_1) . If s cannot reach v via F_0 , the update moves to (v, F_1) with probability one, and since the current state is in Ω it must already be (v, F_1) , so no transition between distinct states occurs. Otherwise both (v, F_0) and (v, F_1) are valid states of the chain, unless $\mathcal{D}_{\mathbf{p}}(F_0) = 0$, in which case $p_e = 1$ and the transition to (v, F_0) has probability $1 - p_e = 0$. By the definition of K_2 ,

$$K_2((v, F_0), (v, F_1)) = \frac{p_e}{m}, \quad K_2((v, F_1), (v, F_0)) = \frac{1 - p_e}{m},$$

and detailed balance for K_2 holds because $\frac{\pi(v, F_1)}{\pi(v, F_0)} = \frac{\mathcal{D}_{\mathbf{p}}(F_1)}{\mathcal{D}_{\mathbf{p}}(F_0)} = \frac{p_e}{1 - p_e}$. □

3.3 Lower bound on the spectral gap

The Markov chain $K = \frac{1}{2}I + \frac{1}{4}K_1 + \frac{1}{4}K_2$ of [Section 2.2](#) is lazy, and by [Proposition 10](#) it is irreducible and reversible with respect to $\pi = \pi_{c, \mathbf{p}}$, so [Lemma 12](#) applies to it. The main technical result of this section is the following lower bound on the spectral gap.

Theorem 15. *For the Markov chain $K = \frac{1}{2}I + \frac{1}{4}K_1 + \frac{1}{4}K_2$, where the vertex weights satisfy $\frac{1}{4q_v} \leq c_v \leq \frac{4}{q_v}$ for all $v \in V$, we have $\text{gap}(K) = \Omega\left(\frac{1}{nm^2}\right)$.*

Note that [Theorem 11](#) is a direct consequence of [Lemma 12](#) and [Theorem 15](#).

3.3.1 Reduction to the ideal chain

The weights $\mathbf{c}$ in [Theorem 15](#) are only assumed to lie within a constant factor of the ideal weights $c_v^* = \frac{1}{q_v}$. The first step is to reduce the general case to the ideal case. Let $\mathbf{c}^* = (c_v^*)_{v \in V}$, let $\pi^* = \pi_{\mathbf{c}^*, \mathbf{p}}$, and let $K^* = \frac{1}{2}I + \frac{1}{4}K_1^* + \frac{1}{4}K_2^*$ be the Markov chain of [Section 2.2](#) with $\mathbf{c}$ replaced by $\mathbf{c}^*$, called the *ideal chain*. Since $\sum_{F \in \Omega_v} \mathcal{D}_{\mathbf{p}}(F) = q_v$, the normalising constant of π^* is $\sum_{v \in V} c_v^* q_v = n$, so that the ideal stationary distribution is

$$\forall (v, F) \in \Omega, \quad \pi^*(v, F) = \frac{\mathcal{D}_{\mathbf{p}}(F)}{n q_v}.$$

In other words, under π^* the marked vertex is uniform over V , and conditioned on the marked vertex being v , the arc set is distributed as $\mathcal{D}_{\mathbf{p}}$ conditioned on Ω_v . The next lemma shows that the two chains have comparable spectral gaps, so that only the ideal chain needs to be analysed.

Lemma 16. *Let $\mathbf{c}$ be as in [Theorem 15](#). It holds that $\text{gap}(K) \geq \frac{1}{4096} \text{gap}(K^*)$.*

Proof. The assumption on the weights says exactly that $c_v q_v \in [\frac{1}{4}, 4]$ for every $v \in V$, so the normalising constant $Z = \sum_{u \in V} c_u q_u$ of π satisfies $Z \in [\frac{n}{4}, 4n]$. For every $(v, F) \in \Omega$,

$$\frac{\pi(v, F)}{\pi^*(v, F)} = \frac{c_v \mathcal{D}_{\mathbf{p}}(F)/Z}{\mathcal{D}_{\mathbf{p}}(F)/(n q_v)} = \frac{n c_v q_v}{Z} \in \left[\frac{n/4}{4n}, \frac{4n}{n/4} \right] = \left[\frac{1}{16}, 16 \right]. \quad (11)$$

We next compare the Dirichlet forms of the two chains. The heat-bath update does not involve the vertex weight parameters, so $K_2^* = K_2$, and (11) gives

$$\pi(X)K_2(X, Y) = \frac{\pi(X)}{\pi^*(X)} \pi^*(X)K_2^*(X, Y) \geq \frac{1}{16} \pi^*(X)K_2^*(X, Y).$$

The Metropolis update uses the same proposal kernel R for both vertex weight vectors, so (10), which holds for any choice of the weights, applies to K_1 and to K_1^* alike. Combining it with (11),

$$\begin{aligned} \pi(X)K_1(X, Y) &= R(X, Y) \min\{\pi(X), \pi(Y)\} \\ &\geq \frac{1}{16} R(X, Y) \min\{\pi^*(X), \pi^*(Y)\} = \frac{1}{16} \pi^*(X)K_1^*(X, Y). \end{aligned}$$

As $\pi(X)K(X, Y) = \frac{1}{4}\pi(X)K_1(X, Y) + \frac{1}{4}\pi(X)K_2(X, Y)$ for $X \neq Y$, and the Dirichlet form involves only such pairs, summing the two bounds yields

$$\forall f: \Omega \rightarrow \mathbb{R}, \quad \mathcal{E}_K(f, f) \geq \frac{1}{16} \mathcal{E}_{K^*}(f, f).$$

For the variances, applying (11) to both factors in (8),

$$\begin{aligned} \text{Var}_{\pi}(f) &= \frac{1}{2} \sum_{X, Y \in \Omega} \pi(X)\pi(Y) (f(X) - f(Y))^2 \\ &\leq \frac{256}{2} \sum_{X, Y \in \Omega} \pi^*(X)\pi^*(Y) (f(X) - f(Y))^2 = 256 \text{Var}_{\pi^*}(f). \end{aligned}$$

Both π and π^* are supported on all of Ω , so $\text{Var}_{\pi}(f) = 0$ if and only if f is constant, in which case $\text{Var}_{\pi^*}(f) = 0$. Using the variational characterisation (9) of the spectral gap, we obtain

$$\text{gap}(K) = \inf_{f: \text{Var}_{\pi}(f) \neq 0} \frac{\mathcal{E}_K(f, f)}{\text{Var}_{\pi}(f)} \geq \frac{1}{4096} \inf_{f: \text{Var}_{\pi^*}(f) \neq 0} \frac{\mathcal{E}_{K^*}(f, f)}{\text{Var}_{\pi^*}(f)} = \frac{1}{4096} \text{gap}(K^*). \quad \square$$

3.3.2 Multicommodity flow for the ideal chain

By Lemma 16, it suffices to prove Theorem 15 for the ideal chain. To ease the notation, the weights are fixed to the ideal ones from now on, and the stars are dropped: throughout the rest of this section, we use the weight $c_v = \frac{1}{q_v}$ for all vertices $v \in V$, and $\pi = \pi^*$ and $K = K^*$ denote the corresponding stationary distribution and Markov chain, so that $\pi(v, F) = \frac{\mathcal{D}_p(F)}{n q_v}$ for all $(v, F) \in \Omega$. The goal is then to show that $\text{gap}(K) = \Omega(\frac{1}{nm^2})$.

We now construct a flow of small congestion, which by Lemma 14 lower bounds the spectral gap. This requires a path between every ordered pair of states, and these paths are obtained by routing everything through the states whose marked vertex is the source. Call these states the s -layer, and denote by

$$\Lambda := \{(s, F) \in \Omega\} = \{(s, F) \mid F \in \Omega_s, \mathcal{D}_p(F) > 0\}.$$

As $q_s = 1$, the ideal distribution satisfies $\pi(s, F) = \frac{\mathcal{D}_p(F)}{n}$ on Λ , so $\pi(\Lambda) = \frac{1}{n}$ and the law of π conditioned on Λ is $\mathcal{D}_p$. It therefore suffices to construct one path from each state in Ω to each state of the s -layer Λ : a path between an arbitrary pair of states is then obtained by concatenating two of them, one taken backwards.

Formally, for every $x \in \Omega$ and every $(s, F') \in \Lambda$, we will construct one path $\gamma_{x, F'}$ in the state graph of K . Given all the paths $(\gamma_{x, F'})_{x \in \Omega, (s, F') \in \Lambda}$, we can construct a set of paths $\mathcal{R}_{x, y}$ from x to y for every ordered pair of distinct states $x, y \in \Omega$ in the following way

$$\mathcal{R}_{x, y} = \left\{ \gamma_{x, y, F'} = \gamma_{x, F'} \circ \gamma_{y, F'}^{-1} \mid (s, F') \in \Lambda \right\},$$

where $\gamma_{y, F'}^{-1}$ is the reversal of $\gamma_{y, F'}$, $\circ$ denotes an operation that first do path concatenation and then erase the cycles of the resulting walk so that $\gamma_{x, y, F'}$ is again a path. We remark that reversibility of the Markov chain makes the state graph symmetric, so the reversal of a path is again a path in the state graph. Define the set of all these paths as $\mathcal{R} = \bigcup_{x \neq y} \mathcal{R}_{x, y}$.

Consider the paths $(\gamma_{x, F'})_{x \in \Omega, (s, F') \in \Lambda}$. For any transition η , define the *load* of η as

$$L(\eta) = \sum_{x \in \Omega} \sum_{(s, F') \in \Lambda} \pi(x) \mathcal{D}_p(F') \cdot \mathbf{1}[\eta \in \gamma_{x, F'}]. \quad (12)$$

Intuitively, $L(\eta)$ is the amount of demand crossing η when every state x sends $\pi(x) \mathcal{D}_p(F')$ units of demand to (s, F') along $\gamma_{x, F'}$, for every $(s, F') \in \Lambda$. The next lemma shows that if we can construct paths $(\gamma_{x, F'})_{x \in \Omega, (s, F') \in \Lambda}$ such that the length of every path is small and the load of every transition is small, then we can define a flow in the sense of Definition 13 that is supported on $\mathcal{R}$ and has small congestion, and as a consequence we can lower bound the spectral gap of K . Recall that for any transition $\eta = (z, z')$, the capacity is $Q(\eta) = \pi(z)K(z, z')$.

Lemma 17. *Suppose that $|\gamma_{x, F'}| \leq \ell$ for every $x \in \Omega$ and every $(s, F') \in \Lambda$, and that $L(\eta) \leq B \cdot Q(\eta)$ for every transition η . Then $\text{gap}(K) \geq \frac{1}{4\ell B}$.*

Proof. Let Γ assign the amount $\pi(x)\pi(y)\mathcal{D}_p(F')$ to the path $\gamma_{x, y, F'}$, for every ordered pair $x \neq y$ and every $(s, F') \in \Lambda$, the contributions of different intermediate states leading to the same path being added up, and let Γ vanish on every path outside $\mathcal{R}$. Since $\sum_{(s, F') \in \Lambda} \mathcal{D}_p(F') = 1$, this is indeed a flow:

$$\forall x \neq y \in \Omega, \quad \sum_{\gamma \in \mathcal{R}_{x, y}} \Gamma(\gamma) = \sum_{(s, F') \in \Lambda} \pi(x)\pi(y)\mathcal{D}_p(F') = \pi(x)\pi(y).$$

It remains to bound the congestion of Γ . Fix a transition $\eta = (z, z')$ and let $\bar{\eta} = (z', z)$ be its reversal, which is also a transition and satisfies $Q(\bar{\eta}) = Q(\eta)$ by reversibility. Erasing cycles only shortens a walk and removes transitions from it, so $|\gamma_{x,y,F'}| \leq 2\ell$, and $\gamma_{x,y,F'}$ traverses η only if $\gamma_{x,F'}$ traverses η or $\gamma_{y,F'}$ traverses $\bar{\eta}$. Therefore

$$\begin{aligned} \sum_{\gamma \in \mathcal{R}: \eta \in \gamma} \Gamma(\gamma) |\gamma| &\leq 2\ell \sum_{x \neq y} \sum_{(s, F') \in \Lambda} \pi(x) \pi(y) \mathcal{D}_p(F') (\mathbf{1}[\eta \in \gamma_{x,F'}] + \mathbf{1}[\bar{\eta} \in \gamma_{y,F'}]) \\ &\leq 2\ell (L(\eta) + L(\bar{\eta})) \leq 4\ell B \cdot Q(\eta), \end{aligned}$$

where the second inequality follows by summing over y in the first term and over x in the second, both of which contribute a factor $\sum_{w \in \Omega} \pi(w) = 1$. Hence $\rho(\Gamma) \leq 4\ell B$, and Lemma 14 completes the proof. $\square$

Next, we construct the paths $(\gamma_{x,F'})_{x \in \Omega, (s, F') \in \Lambda}$. Fix $x = (u, F) \in \Omega$ and $(s, F') \in \Lambda$. Fix a total order of the arcs in E . Run a modified breadth-first search (BFS) from s in G , stopping as soon as u is discovered. This process is essentially a BFS on (V, F) from s , except that we want to record the order of arcs being inspected. To be more precise, we maintain a first-in-first-out (FIFO) queue of discovered but not yet explored vertices. When a vertex is removed from the queue, inspect *all* of its outgoing arcs $e \in E$ from smallest to largest according to the fixed order. If $e \notin F$, skip it. If $e \in F$ and its head w has not yet been discovered, discover w and add it to the queue. The process continues until u is discovered. Record e in the order in which it is inspected, and append all the uninspected arcs from smallest to largest according to the fixed order. This gives a permutation $a_1, a_2, \dots, a_m$ of all arcs in E . As $(u, F) \in \Omega$, the search must discover u . Let

$$s = v_0 \longrightarrow v_1 \longrightarrow \dots \longrightarrow v_r = u$$

be the discovery path from s to u . For every $1 \leq i \leq r$, let τ_i be the position of the discovery arc (v_{i-1}, v_i) in the permutation, so that $a_{\tau_i} = (v_{i-1}, v_i)$. If $r \geq 1$, the BFS construction ensures that $1 \leq \tau_1 < \dots < \tau_r \leq m$. In all cases, set $\tau_0 = 0$ and $\tau_{r+1} = m + 1$. An example is given in Figure 4.

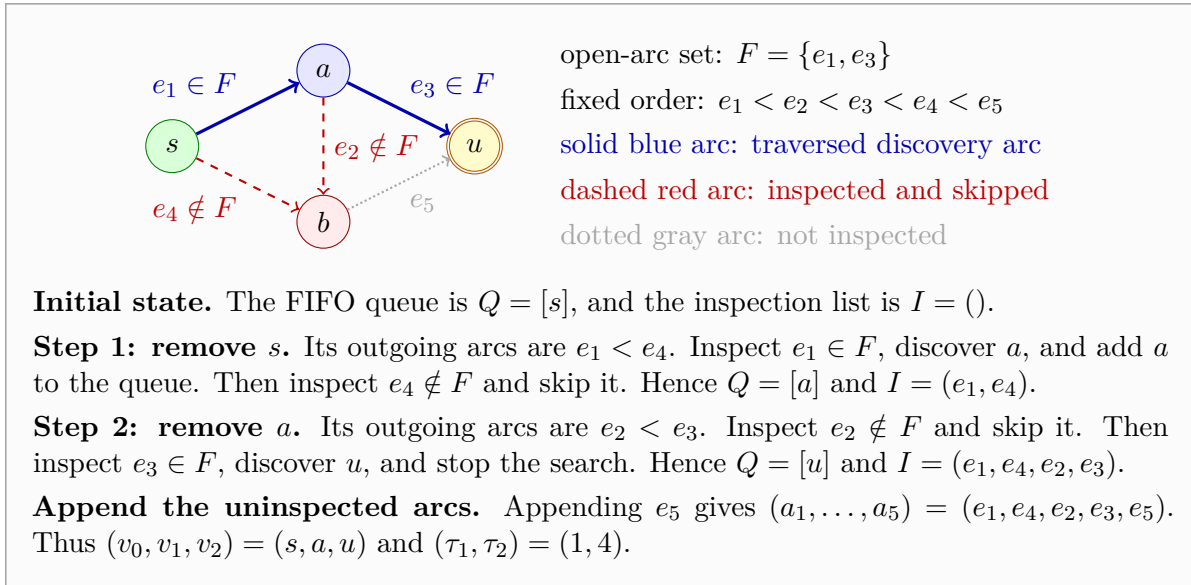


Figure 4: An example of the modified BFS and the resulting arc permutation.

For every $0 \leq k \leq m$, define an arc set $H^k \subseteq E$ by taking its first k arcs in the permutation $a_1, \dots, a_m$ according to F and its remaining arcs according to F' , that is,

$$\mathbf{1}[a_j \in H^k] = \begin{cases} \mathbf{1}[a_j \in F], & j \leq k, \\ \mathbf{1}[a_j \in F'], & j > k, \end{cases} \quad 1 \leq j \leq m.$$

Thus $H^m = F$ and $H^0 = F'$. It is easy to see that every H^k has positive probability under $\mathcal{D}_p$.

We now construct the path $\gamma_{x,F'}$ from $x = (u, F)$ to (s, F') by scanning the arcs in the reverse order $a_m, a_{m-1}, \dots, a_1$, starting from $(v_r, H^m) = (u, F)$. Basically, we flip the arc if it is in $F \oplus F'$, and in the meantime go backwards from u to s along the discovery path. Note that we can move the marked vertex only if the current arc is the discovery arc entering it. Consider the iteration that scans a_k and let $i \in \{0, \dots, r\}$ be the unique index satisfying $\tau_i \leq k < \tau_{i+1}$. The construction maintains the following invariant: at the beginning of the iteration scanning a_k , the current state is (v_i, H^k) . The invariant holds when the scan starts, namely for $k = m$, because then $i = r$ and the starting state is $(v_r, H^m) = (u, F)$. Now suppose the invariant holds at the beginning of the iteration that scans a_k . There are two cases.

- If $k \notin \{\tau_1, \dots, \tau_r\}$, keep the marked vertex v_i fixed and replace H^k by H^{k-1} . If $\mathbf{1}[a_k \in F] = \mathbf{1}[a_k \in F']$, then $H^k = H^{k-1}$ and no transition is needed. Otherwise, this change is a transition of K_2 . Indeed, every arc on the discovery path from s to v_i belongs to both H^k and H^{k-1} , so v_i remains reachable after the update. The resulting state is (v_i, H^{k-1}) .
- If $k = \tau_i$ for some $1 \leq i \leq r$, then $a_k = (v_{i-1}, v_i)$ and $a_k \in H^k \cap F$. Use a K_1 transition to move the marked vertex from v_i to v_{i-1} and at the same time replace H^k by H^{k-1} . If $a_k \in F'$, this is the proposal that retains a_k ; otherwise, it is the proposal that deletes a_k . In either case the proposed state is (v_{i-1}, H^{k-1}) . It is feasible because every arc on the shorter discovery path from s to v_{i-1} still belongs to H^{k-1} , and the Metropolis acceptance probability is positive because both states belong to the support of π .

In both cases the invariant is maintained at the start of the next scan: in the first case $\tau_i < k$, so that $\tau_i \leq k-1 < \tau_{i+1}$ and the iteration ends in the state (v_i, H^{k-1}) , and in the second case $k = \tau_i$, so that $\tau_{i-1} \leq k-1 < \tau_i$ and the iteration ends in the state (v_{i-1}, H^{k-1}) . After the whole reverse scan, the arc set has changed from F to F' , and the marked vertex has moved backwards along the discovery path from $v_r = u$ to $v_0 = s$. The path $\gamma_{x,F'}$ is obtained by keeping only the non-trivial transitions, and the end state is $(v_0, H^0) = (s, F')$. Each scan contributes at most one transition, so

$$|\gamma_{x,F'}| \leq m. \tag{13}$$

Moreover, while the marked vertex is fixed, every nontrivial K_2 transition changes a distinct arc. Every nontrivial K_1 transition moves the marked vertex one step backwards along the simple discovery path, so once the construction leaves a vertex, it never returns to it. Hence no state is visited twice, and $\gamma_{x,F'}$ is a simple path in the transition graph of K .

3.4 Analysis of the load and congestion

The key property of the flow defined in the last subsection is its low load and congestion.

Lemma 18. *The family of paths $(\gamma_{x,F'})_{x \in \Omega, (s,F') \in \Lambda}$ constructed above satisfies*

$$L(\eta) = \sum_{x \in \Omega} \sum_{(s,F') \in \Lambda} \pi(x) \mathcal{D}_p(F') \cdot \mathbf{1}[\eta \in \gamma_{x,F'}] \leq 8nmQ(\eta)$$

for every transition $\eta = (z, z')$ of K .

Proof. Fix an oriented transition $\eta = (z, z')$ and a value $u \in V$ for the marked vertex. Consider all pairs (F, F') such that $(u, F) \in \Omega$, $(s, F') \in \Lambda$, and $\gamma_{(u,F),F'}$ traverses η . Formally, define the set

$$\mathcal{A}_{\eta,u} := \{(F, F') \mid (u, F) \in \Omega, (s, F') \in \Lambda, \text{ and } \gamma_{(u,F),F'} \text{ traverses } \eta\}.$$

If $\mathcal{A}_{\eta,u} = \emptyset$, its contribution is zero. We may therefore assume that $\mathcal{A}_{\eta,u} \neq \emptyset$. Write $z = (w, B)$ and $z' = (w', B')$. The transition η determines a distinguished arc a^* . For a K_2 transition, a^* is the unique arc on which B and B' differ. For a K_1 transition, the construction moves the marked vertex backwards from w to w' , and $a^* = (w', w)$ is its discovery arc; this arc is unique because G is simple.

Fix a pair $(F, F') \in \mathcal{A}_{\eta,u}$. The permutation $(a_1, \dots, a_m)$ is determined by (u, F) and may therefore vary with the pair. Let $k = k(F, F')$ be the unique position of a^* in this permutation, so that $a_k = a^*$. The process of constructing the path $\gamma_{(u,F),F'}$ from $x = (u, F)$ to (s, F') defines a sequence of arc sets $F' = H^0, H^1, \dots, H^m = F$, and it traverses η from $(w, H^k) = (w, B)$ to $(w', H^{k-1}) = (w', B')$ while scanning a^* . At this pair-dependent scan position k , define the encoding $\tilde{H}^k \subseteq E$ by

$$\mathbf{1}[a_j \in \tilde{H}^k] = \begin{cases} \mathbf{1}[a_j \in F'], & j \leq k, \\ \mathbf{1}[a_j \in F], & j > k, \end{cases} \quad 1 \leq j \leq m.$$

The hybrid H^k and the encoding $\tilde{H}^k$ are determined by (F, F') together with the corresponding permutation, and their dependence on the chosen pair is suppressed in the notation. For every arc $e \in E$, the two indicators $\mathbf{1}[e \in H^k]$ and $\mathbf{1}[e \in \tilde{H}^k]$ are $\mathbf{1}[e \in F]$ and $\mathbf{1}[e \in F']$, possibly in the opposite order. In other words, $\tilde{H}^k \oplus H^k = F \oplus F'$ and $\tilde{H}^k \cap H^k = F \cap F'$. Since $\mathcal{D}_p$ is a product measure, we have

$$\mathcal{D}_p(F)\mathcal{D}_p(F') = \mathcal{D}_p(H^k)\mathcal{D}_p(\tilde{H}^k). \quad (14)$$

We now consider the two types of transitions separately. First, suppose that $\eta = (z, z')$ is a nontrivial transition of K_2 . Then $w' = w$, and the arc sets B and B' differ exactly at a^* ; in particular, $B \setminus \{a^*\} = B' \setminus \{a^*\}$. Such a transition can only occur when $p_{a^*} < 1$. Since the heat-bath update chooses a^* with probability $1/m$ and K chooses K_2 with probability $1/4$, its capacity is

$$Q(\eta) = \pi(z)K(z, z') = \frac{\mathcal{D}_p(B)}{nq_w}K(z, z') = \frac{\mathcal{D}_{p,-a^*}(B \setminus \{a^*\})p_{a^*}(1 - p_{a^*})}{4mnq_w}, \quad (15)$$

where $\mathcal{D}_{p,-a^*}$ denotes the product measure on arc subsets of $E \setminus \{a^*\}$.

Recall that (u, F) needs to send $\pi(u, F)\mathcal{D}_p(F')$ units of demand to (s, F') along $\gamma_{(u,F),F'}$. For every $(F, F') \in \mathcal{A}_{\eta,u}$, exactly one of H^k and $\tilde{H}^k$ contains a^* , that is,

$$\{\mathbf{1}[a^* \in H^k], \mathbf{1}[a^* \in \tilde{H}^k]\} = \{0, 1\}.$$

Moreover, $B = H^k$ and $B' = H^{k-1}$. Therefore (14) and (15) give

$$\begin{aligned}\pi(u, F)\mathcal{D}_{\mathbf{p}}(F') &= \frac{\mathcal{D}_{\mathbf{p}}(F)\mathcal{D}_{\mathbf{p}}(F')}{nq_u} = \frac{\mathcal{D}_{\mathbf{p}}(H^k)\mathcal{D}_{\mathbf{p}}(\tilde{H}^k)}{nq_u} \\ &= \frac{p_{a^*}(1-p_{a^*})}{nq_u} \mathcal{D}_{\mathbf{p}, -a^*}(B \setminus \{a^*\}) \mathcal{D}_{\mathbf{p}, -a^*}(\tilde{H}^k \setminus \{a^*\}) \\ &= 4mQ(\eta) \cdot \frac{q_w}{q_u} \mathcal{D}_{\mathbf{p}, -a^*}(\tilde{H}^k \setminus \{a^*\}).\end{aligned}$$

To upper bound the overall contribution from $\mathcal{A}_{\eta, u}$, we use the following claim. The proof is deferred until after both types of transitions have been analysed.

Claim 19. *For the fixed η and u , where $k = k(F, F')$ in each summand,*

$$\sum_{(F, F') \in \mathcal{A}_{\eta, u}} \frac{q_w}{q_u} \mathcal{D}_{\mathbf{p}, -a^*}(\tilde{H}^k \setminus \{a^*\}) \leq 1. \quad (16)$$

Summing first over all demands with fixed u , applying Claim 19 to the resulting inner sum, and then summing over all choices of u , we obtain

$$\begin{aligned}L(\eta) &= \sum_{u \in V} \sum_{(F, F') \in \mathcal{A}_{\eta, u}} \pi(u, F)\mathcal{D}_{\mathbf{p}}(F') \\ &= 4mQ(\eta) \sum_{u \in V} \sum_{(F, F') \in \mathcal{A}_{\eta, u}} \frac{q_w}{q_u} \mathcal{D}_{\mathbf{p}, -a^*}(\tilde{H}^k \setminus \{a^*\}) \leq 4nmQ(\eta).\end{aligned}$$

It remains to consider transitions η of K_1 . The transition $\eta = (z, z')$ moves the marked vertex backwards from w to w' while scanning the discovery arc $a^* = (w', w)$. Since a^* is a discovery arc, $a^* \in B$. Moreover, B and B' differ at most at a^* . Hence $B \setminus \{a^*\} = B' \setminus \{a^*\}$. Define

$$\theta = \theta_{a^*}(B') := \begin{cases} p_{a^*}, & a^* \in B', \\ 1 - p_{a^*}, & a^* \notin B'. \end{cases}$$

Selecting the arc a^* and the corresponding one of the two proposals in K_1 has probability $1/(2m)$. Other proposals may lead to the same proposed state z' , but they can only increase the capacity.² Hence the detailed balance condition (10) and the coefficient $1/4$ of K_1 in K imply

$$\begin{aligned}Q(\eta) &\geq \frac{1}{8m} \min\{\pi(z), \pi(z')\} = \frac{1}{8mn} \min\left\{\frac{\mathcal{D}_{\mathbf{p}}(B)}{q_w}, \frac{\mathcal{D}_{\mathbf{p}}(B')}{q_{w'}}\right\} \\ (\text{by } B \setminus \{a^*\} &= B' \setminus \{a^*\}) = \frac{\mathcal{D}_{\mathbf{p}, -a^*}(B \setminus \{a^*\})}{8mn} \min\left\{\frac{p_{a^*}}{q_w}, \frac{\theta}{q_{w'}}\right\}.\end{aligned} \quad (17)$$

We will use the following inequality:

$$q_w \geq p_{a^*} q_{w'}. \quad (18)$$

Indeed, the event that s can reach w' is independent of the outgoing arc $a^* = (w', w)$: whenever s can reach w' , there is a simple path witnessing this event, and such a path does not use a^* . Thus,

²This can occur when $a^* \in B'$ and the reverse arc (w, w') also belongs to B .

with probability $p_{a^*}q_{w'}$, the vertex s can reach w' and the arc a^* is open. When this happens, s can reach w . Since $\theta \leq 1$, and (18) gives $\theta/q_{w'} \geq p_{a^*}\theta/q_w$, both terms in the minimum in (17) are at least $p_{a^*}\theta/q_w$. Consequently,

$$Q(\eta) \geq \frac{\mathcal{D}_{\mathbf{p}, -a^*}(B \setminus \{a^*\})p_{a^*}\theta}{8mnq_w}. \quad (19)$$

For every $(F, F') \in \mathcal{A}_{\eta, u}$, we have $H^k = B$ and $H^{k-1} = B'$. Hence $a^* \in H^k$ and $\mathbf{1}[a^* \in \tilde{H}^k] = \mathbf{1}[a^* \in F'] = \mathbf{1}[a^* \in B']$. Applying (14) and (19),

$$\begin{aligned} \pi(u, F)\mathcal{D}_{\mathbf{p}}(F') &= \frac{\mathcal{D}_{\mathbf{p}}(F)\mathcal{D}_{\mathbf{p}}(F')}{nq_u} = \frac{\mathcal{D}_{\mathbf{p}}(H^k)\mathcal{D}_{\mathbf{p}}(\tilde{H}^k)}{nq_u} \\ &= \frac{\mathcal{D}_{\mathbf{p}, -a^*}(B \setminus \{a^*\})p_{a^*}\theta\mathcal{D}_{\mathbf{p}, -a^*}(\tilde{H}^k \setminus \{a^*\})}{nq_u} \leq 8mQ(\eta) \cdot \frac{q_w}{q_u}\mathcal{D}_{\mathbf{p}, -a^*}(\tilde{H}^k \setminus \{a^*\}). \end{aligned}$$

As in the K_2 case, summing first over all demands with fixed u , applying Claim 19 to the resulting inner sum, and then summing over all choices of u , we obtain

$$L(\eta) = \sum_{u \in V} \sum_{(F, F') \in \mathcal{A}_{\eta, u}} \pi(u, F)\mathcal{D}_{\mathbf{p}}(F') \leq 8nmQ(\eta).$$

This completes the proof of Lemma 18, subject to Claim 19. $\square$

We then prove Claim 19. We will use the following classical positive-correlation inequality for product measures. Let I be a finite set and, for each $i \in I$, let μ_i be a probability measure on $\{0, 1\}$. Let $\mu = \bigotimes_{i \in I} \mu_i$ be the product measure on $\{0, 1\}^I$. On the space $\{0, 1\}^I$, define a partial order $x \preceq y$ if $x_i \leq y_i$ for every $i \in I$. We say that an event $A \subseteq \{0, 1\}^I$ is increasing if $x \in A$ and $x \preceq y$ implies $y \in A$.

Lemma 20 (Harris's inequality [Har60]). *If $\mathcal{A}$ and $\mathcal{B}$ are increasing events, $\mu(\mathcal{A} \cap \mathcal{B}) \geq \mu(\mathcal{A})\mu(\mathcal{B})$.*

Proof of Claim 19. For each pair $(F, F') \in \mathcal{A}_{\eta, u}$, let $k = k(F, F')$ be as above and call the arc set $\tilde{H}^k \setminus \{a^*\} \subseteq E \setminus \{a^*\}$ its *encoding*. We first show that the encoding map is injective even though k is not fixed, namely that (F, F') can be uniquely recovered from the encoding $\tilde{H}^k \setminus \{a^*\}$ even if the value of k is not known. The transition η determines the arc sets $H^k = B$ and $H^{k-1} = B'$ before and after the transition, and hence also determines whether a^* belongs to $\tilde{H}^k$, because

$$\mathbf{1}[a^* \in \tilde{H}^k] = \mathbf{1}[a^* \in F'] = \mathbf{1}[a^* \in H^{k-1}].$$

Thus the encoding, together with η , determines both H^k and the entire arc set $\tilde{H}^k$. It remains to recover the permutation and the position k . Recall that constructing the permutation requires running the modified BFS from s until u is discovered. We replay this BFS without knowing k , using H^k to determine membership in F until and including the inspection of a^* , and using $\tilde{H}^k$ strictly after a^* has been inspected.

More precisely, suppose that the first $\ell - 1$ inspected arcs have been recovered. These arcs and their recovered inspection outcomes determine the full BFS control state—the queue, the discovered vertices, the vertex currently being explored, and its outgoing-arc cursor—and hence determine the next inspected arc a_ℓ . If a^* has not yet been inspected, use $\mathbf{1}[a_\ell \in H^k]$ as the

indicator of membership in F , and switch to $\tilde{H}^k$ only after processing a_ℓ when $a_\ell = a^*$. If a^* has already been inspected, use $\mathbf{1}[a_\ell \in \tilde{H}^k]$. In either case this indicator determines the outcome of the inspection, and hence the next BFS control state. Inductively, this recovers the inspected arcs in their original order until u is discovered.

If a^* has been inspected by that time, its position in the recovered sequence is k . Otherwise, the BFS has stopped before inspecting a^* ; append the remaining uninspected arcs in the fixed order, as in the original construction, and locate a^* in this appended suffix. This again recovers both the entire permutation $a_1, \dots, a_m$ and the position k . Finally, for positions at most k , the indicators of F and F' are respectively those of H^k and $\tilde{H}^k$, while after position k their roles are reversed. Hence both F and F' are recovered, proving that the encoding map is injective.

We have also recovered the discovery path from s to u . Recall that $\eta = (z, z')$ is fixed and that w is the marked vertex of z ; in particular, w is determined by η . Write $w = v_i$ on the recovered discovery path. The invariant in the path construction gives $\tau_i \leq k < \tau_{i+1}$ (and $k = \tau_i$ in the K_1 case). Therefore, every arc on the suffix

$$w = v_i \longrightarrow v_{i+1} \longrightarrow \dots \longrightarrow v_r = u$$

of the discovery path has position greater than k . Hence every such arc belongs to $\tilde{H}^k$ and none of them is the arc a^* . We obtain the following key observation.

Observation 21. *The vertex w can reach u via arcs in $\tilde{H}^k \setminus \{a^*\}$.*

The injectivity proved above shows that distinct pairs $(F, F') \in \mathcal{A}_{\eta, u}$ give distinct encodings. Moreover, by [Observation 21](#), each such encoding contains a path from w to u . Since this path does not use a^* , we obtain

$$\begin{aligned} \sum_{(F, F') \in \mathcal{A}_{\eta, u}} \mathcal{D}_{p, -a^*}(\tilde{H}^k \setminus \{a^*\}) &\leq \Pr_{\mathcal{D}_p} [w \text{ can reach } u \text{ without using } a^*] \\ &\leq \Pr_{\mathcal{D}_p} [w \text{ can reach } u]. \end{aligned} \tag{20}$$

To apply [Lemma 20](#), identify each arc set $S \subseteq E$ with its indicator vector in $\{0, 1\}^E$. Under this identification, $\mathcal{D}_p = \bigotimes_{e \in E} \text{Bernoulli}(p_e)$ is a product measure, and the coordinatewise partial order is precisely set inclusion of arc sets. The events that s can reach w and that w can reach u are increasing: if an arc set contains a path witnessing either event, then every larger arc set contains the same path. Applying [Lemma 20](#) to these two events, and observing that their intersection implies that s can reach u , gives

$$\begin{aligned} q_u &\geq \Pr_{\mathcal{D}_p} [s \text{ can reach } w \text{ and } w \text{ can reach } u] \\ &\geq \Pr_{\mathcal{D}_p} [s \text{ can reach } w] \Pr_{\mathcal{D}_p} [w \text{ can reach } u] = q_w \Pr_{\mathcal{D}_p} [w \text{ can reach } u]. \end{aligned} \tag{21}$$

Combining (20) and (21) proves [Claim 19](#) and completes the proof. $\square$

3.4.1 Proof of the lower bound on the spectral gap

Finally, we put all the pieces together to prove the lower bound on the spectral gap in [Theorem 15](#).

Proof of Theorem 15. First consider the ideal weights $c_v^* = 1/q_v$; recall that in the preceding construction this ideal chain was denoted by K after dropping the stars. By (13), the paths constructed

above have length at most $\ell = m$, and by [Lemma 18](#) their load satisfies $L(\eta) \leq 8nmQ(\eta)$ for every transition η . Thus we may take $B = 8nm$ in [Lemma 17](#), which gives

$$\text{gap}(K^*) \geq \frac{1}{4\ell B} \geq \frac{1}{4m \cdot 8nm} = \frac{1}{32nm^2}.$$

For the original weights $\mathbf{c}$, [Lemma 16](#) now yields

$$\text{gap}(K) \geq \frac{1}{4096} \text{gap}(K^*) = \Omega\left(\frac{1}{nm^2}\right). \quad \square$$

3.5 Lower bounds and discussion on the mixing time

Now we discuss some possible improvement and limitations of the Markov chain. First, to avoid the $\Omega(n^3)$ lower bound on paths (or multiple paths between s and t), we can change K_1 to propose an adjacent arc from the marked vertex instead of a uniformly at random arc, and then use the Metropolis filter to get the correct stationary distribution. This change would improve the mixing time to $O(n^2)$ on paths, but it would not improve the overall upper bound in the mixing time analysis above. It is because the contribution in the congestion from K_1 and K_2 in the current analysis are of the same order, and improving the contribution from K_1 alone would not change the final bound. For simplicity of presentation, we choose to use the random arc proposal version of K_1 .

On the other hand, even for the adjacent proposal version of K_1 , there is an $\Omega(n^3)$ lower bound on the inverse spectral gap. Here is an example. Fix p to be the open probability for all edges. Let r be an odd number, and $s = v_0, v_1, \dots, v_r$ be a directed path from s to v_r , and introduce new vertices $u_1, \dots, u_r$. Add arcs (v_i, u_j) for $0 \leq i \leq r$ and $1 \leq j \leq r$ unless $i = j \neq r$. Moreover, add a set of arcs from a matching $(u_1, u_2), (u_3, u_4), \dots, (u_{r-2}, u_{r-1})$. This way, all vertices have the same number of adjacent arcs, which is not necessary for the lower bound but makes the calculation easier. Moreover, $n = 2r + 1$, and $m = \Theta(n^2)$. To show the lower bound on the inverse spectral gap, we use [\(9\)](#) and let the testing function f be defined by $f(v_i, F) = i$ for any F and $0 \leq i \leq r$, and $f(u_i, F) = 0$ for any F and $1 \leq i \leq r$. Since I and K_2 do not change the marked vertex, they do not contribute anything to the Dirichlet form, and only the contribution from K_1 matters. We omit the calculation details, and the conclusion here is that $\text{Var}_\pi(f) = \Theta(n^2)$ and $\mathcal{E}_K(f, f) = O_p(n^{-1})$, which implies the $\Omega(n^3)$ lower bound on the inverse spectral gap. Essentially, this is still the path example, except that a lot of arcs are added to reduce the useful proposal probability of arcs in K_1 .

In any case, our mixing time bound is still a polynomial factor away from the lower bound above, especially in dense graphs. To further improve the mixing time analysis, one may need the local coupling technique from [\[CFJ⁺25\]](#). However that appears to be non-trivial and we leave that for future work.

4 Remaining proofs for FPRAS

In this section, we give the missing proofs for the FPRAS, namely Item [\(2\)](#) of [Lemma 5](#) and the proof of [Lemma 8](#). Both proofs are standard and straightforward and we give them here for completeness.

Proof of Item (2) of Lemma 5. Abbreviate $q_v = q_v(\mathbf{p})$ and $\pi = \pi_{\mathbf{c}, \mathbf{p}}$, and let

$$Z = \sum_{u \in V} c_u q_u \quad \text{and} \quad r_v = \sum_{F \in \Omega_v} \pi(v, F) = \frac{c_v q_v}{Z}.$$

Thus, r_v is the probability that the marked vertex of a sample from π equals v . The hypothesis on the weights says that $c_v q_v \in [\frac{1}{4}, 4]$ for every $v \in V$, so $Z \in [\frac{n}{4}, 4n]$ and hence

$$r_v \geq \frac{1}{16n} \quad \forall v \in V. \quad (22)$$

Define the number of samples as

$$N := \left\lceil \frac{768n}{\varepsilon^2} \log \left(\frac{8n}{\delta} \right) \right\rceil = O \left(\frac{n}{\varepsilon^2} \log \left(\frac{n}{\delta} \right) \right).$$

Let $X_1, \dots, X_N$ be independent ideal samples from π , and let N_v be the number of these samples whose marked vertex equals v . Then $\mathbb{E}[N_v] = Nr_v$. By the multiplicative Chernoff bound and (22), by setting $\alpha = \frac{\varepsilon}{4}$, we have

$$\Pr[|N_v - Nr_v| > \alpha Nr_v] \leq 2 \exp \left(-\frac{\alpha^2 Nr_v}{3} \right) \leq 2 \exp \left(-\frac{\varepsilon^2 N}{768n} \right) \leq \frac{\delta}{4n}.$$

Therefore, with probability at least $1 - \frac{\delta}{4}$, we have $(1 - \alpha)Nr_v \leq N_v \leq (1 + \alpha)Nr_v$ for every $v \in V$. Since $q_s = 1$, the definition of r_v gives, for the ideal-sample ratio estimator $\bar{q}_v = \frac{c_s}{c_v} \frac{N_v}{N_s}$,

$$\frac{\bar{q}_v}{q_v} = \frac{c_s}{c_v q_v} \cdot \frac{N_v}{N_s} = \frac{r_s}{r_v} \cdot \frac{N_v}{N_s} = \frac{N_v / (Nr_v)}{N_s / (Nr_s)} \in \left[\frac{1 - \alpha}{1 + \alpha}, \frac{1 + \alpha}{1 - \alpha} \right] \subseteq [1 - \varepsilon, 1 + \varepsilon], \quad \forall v \in V,$$

where the last inclusion follows from $\alpha = \frac{\varepsilon}{4}$ and $\varepsilon < 1$.

We now replace the ideal samples by actual ones. Run N independent copies of the Markov chain in Item (1) of Lemma 5. For every copy j , independently draw $F_{0,j} \sim \mathcal{D}_{\mathbf{p}}$, start the chain from $(s, F_{0,j})$, use independent transition randomness, and run it with total variation error bounded by $\eta = \frac{\delta}{4N}$. Denote the outputs of these uncapped runs by $\tilde{X}_1, \dots, \tilde{X}_N$, and let $\tilde{N}_v$ be the number of these outputs whose marked vertex is v . By the coupling characterisation of the total variation distance and a union bound, the ideal and actual samples can be coupled identically with probability at least $1 - N\eta = 1 - \frac{\delta}{4}$.

It remains to impose a worst-case running-time bound. Let B be the bound guaranteed by Item (1) of Lemma 5 with $\eta = \frac{\delta}{4N}$ and $\zeta = \frac{\delta}{2N}$, so that $B = O \left(n^2 m^2 \log \left(\frac{4nN}{\delta} \right) + m \log \left(\frac{2N}{\delta} \right) \right)$. Each uncapped run takes at most B time except with probability at most $\frac{\delta}{2N}$. Hence, by a union bound, the total running time of all N uncapped runs is at most NB with probability at least $1 - \frac{\delta}{2}$. The algorithm simulates the runs sequentially while maintaining an operation counter, including during every breadth-first search, and stops before executing any operation that would cause the count to exceed NB , in which case it returns the all-one vector. If all runs finish before the cutoff but $\tilde{N}_v = 0$ for some $v \in V$, it also returns the all-one vector. Otherwise, it returns

$$\hat{q}_v = \frac{c_s}{c_v} \frac{\tilde{N}_v}{\tilde{N}_s} \quad (v \in V).$$

On the intersection of the concentration and coupling events above, $\tilde{N}_v = N_v > 0$ for every v , and the returned estimates equal the ideal-sample estimates $\bar{q}_v$. Thus the zero-count fallback introduces no additional failure event. The operation cutoff bounds the running time in the worst case and changes the output only on an event of probability at most $\frac{\delta}{2}$. A union bound over the concentration, coupling, and cutoff events proves the claimed approximation guarantee with probability at least $1 - \frac{\delta}{4} - \frac{\delta}{4} - \frac{\delta}{2} = 1 - \delta$.

Substituting the choice of N into NB gives

$$O\left(\frac{n^3 m^2}{\varepsilon^2} \log\left(\frac{n}{\varepsilon \delta}\right) \log\left(\frac{n}{\delta}\right)\right).$$

The additional time needed to form the counts and compute the estimates is absorbed in this bound, which completes the proof. $\square$

Proof of Lemma 8. Set the length of the schedule as

$$L = \min\left\{j \geq 1 : \beta^j \leq p_{\min}\right\} = O\left(n \log\left(\frac{2}{p_{\min}}\right)\right).$$

Indeed, since $\log(1/\beta) = -\log(1 - 1/(4n)) \geq 1/(4n)$,

$$L \leq 1 + \frac{\log(1/p_{\min})}{\log(1/\beta)} \leq 1 + 4n \log\left(\frac{1}{p_{\min}}\right) = O\left(n \log\left(\frac{2}{p_{\min}}\right)\right).$$

For every $e \in E$ and $0 \leq i \leq L$, define $p_e^{(i)} = \max\{p_e, \beta^i\}$. We verify that this construction satisfies Definition 7. By the definition of L ,

$$p_e^{(0)} = \max\{p_e, 1\} = 1, \quad p_e^{(L)} = \max\{p_e, \beta^L\} = p_e.$$

Thus it satisfies the endpoints $\mathbf{p}^{(0)} = \mathbf{1}$ and $\mathbf{p}^{(L)} = \mathbf{p}$.

For every $e \in E$ and $1 \leq i \leq L$,

$$\begin{aligned} \beta p_e^{(i-1)} &= \beta \max\{p_e, \beta^{i-1}\} = \max\{\beta p_e, \beta^i\} \\ &\leq \max\{p_e, \beta^i\} = p_e^{(i)} \leq p_e^{(i-1)}. \end{aligned}$$

Hence the sequence satisfies Definition 7.

Finally, the thresholds can be computed recursively, and constructing each probability vector requires one scan over the m arcs. The total running time is $O(m(L+1)) = O(mL)$. $\square$

5 Hardness for unreliability

At last, we show complementary hardness for the $s-t$ unreliability, Theorem 3. Let us define some notation first. Let $G = (V, E)$ be an undirected graph. We consider the case where all edges open with the same probability independently. This only makes our hardness result stronger. Let G_p be the random subgraph where each edge $e \in E$ is open with probability p independently. Then the $s-t$ unreliability is defined as follows, where $s, t \in V$:

$$\text{UnRel}_G(s, t; p) := \Pr_{G_p}[s \not\rightarrow t] = \sum_{\substack{F \subseteq E: \\ s \text{ cannot reach } t \text{ in } (V, F)}} p^{|F|} (1-p)^{|E \setminus F|}.$$

Now we are ready to prove [Theorem 3](#). We note that the main construction here is due to Provan and Ball [[PB83](#)].

Proof of Theorem 3. Given a bipartite graph $G = (L \cup R, E)$, where L and R are the two vertex classes. We may assume that G contains no isolated vertices. We first construct a new graph $H = (V, E_H)$ as follows. The vertex set is $V = L \cup R \cup \{s, t\}$. For each edge $\{u, v\} \in E$, where $u \in L$ and $v \in R$, add one copy of each of the edges $\{s, u\}$, $\{u, v\}$, and $\{v, t\}$ to E_H . Note that this will introduce parallel edges, and $|E_H| = 3|E|$. An illustration is given in [Figure 5](#).

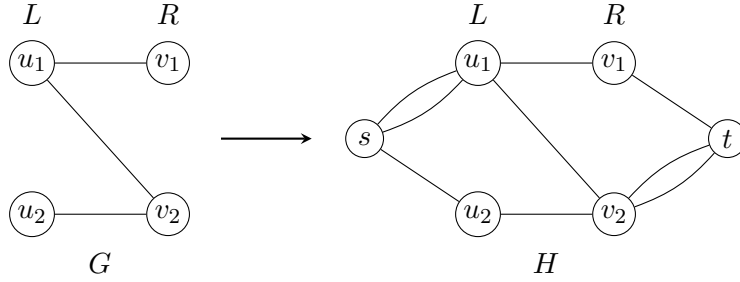


Figure 5: The transformation from G to H . Each edge $\{u, v\} \in E$, with $u \in L$ and $v \in R$, contributes one copy of each of $\{s, u\}$, $\{u, v\}$, and $\{v, t\}$.

We call a subset S of vertices an s -set if $s \in S$ and $t \notin S$. Let $\delta_H(S)$ denote the cut set induced by S in H , and let $I(S) = (L \setminus S) \cup (S \cap R)$. We claim that

$$|\delta_H(S)| = |E| + 2e_H(I(S)), \quad (23)$$

where $e_H(I(S))$ denotes the number of edges in H between vertices of $I(S)$. Note that by definition, $I(S) \subseteq L \cup R$ and hence $e_H(I(S)) = e_G(I(S))$. To see the claim, consider each edge $\{u, v\} \in E$. It contributes 3 to the cut if both u and v are in $I(S)$, and contributes 1 otherwise. Thus, by (23), S defines a minimum $s-t$ cut in H if and only if $I(S)$ is an independent set in G , and this is a bijection.

Now, suppose the number of independent sets in G is Z , and the open probability is $q > 0$. Let $n := |L \cup R|$ and $m := |E|$. We have that

$$Z(1-q)^m q^{2m} \leq \text{UnRel}_H(s, t; q) \leq Z(1-q)^m + 2^n (1-q)^{m+2}. \quad (24)$$

For each $F \subseteq E_H$ such that s cannot reach t in (V, F) , it naturally induces an s -set S_F , which is the set of vertices reachable from s . For any s -set S that can be induced this way, the total probability of all F such that $S_F = S$ is at least $(1-q)^{|\delta_H(S)|} q^{3m - |\delta_H(S)|}$ and at most $(1-q)^{|\delta_H(S)|}$. Then the lower bound follows from the aforementioned bijection. For the upper bound, we use (23), and the two terms correspond to minimum $s-t$ cuts and other $s-t$ cuts, respectively. Thus, the estimates (24) imply that if $1-q = \exp(-Cn/\varepsilon)$ for a sufficiently large constant C , $\frac{\text{UnRel}_H(s, t; q)}{(1-q)^m}$ is a $1 \pm \varepsilon$ approximation of Z .

The remaining question is how to implement exponentially small close probability. This is easy because we can replace each edge in H by ℓ parallel paths of length 2. If each new edge is open with probability p , then effectively, this implements a close probability of $(1 - p^2)^\ell$, and also removes the use of parallel edges. For our needs, it suffices to set $\ell := O(n/\varepsilon)$. The resulting graph has size polynomial in n and $1/\varepsilon$, which finishes the reduction. $\square$

Acknowledgments and use of AI

Weiming Feng thanks Tianyi Zhang for extensive and insightful discussions on the two-terminal reliability problem during his time at ETH Zürich in 2024. He also acknowledges support from the Hong Kong Research Grants Council under ECS Grant No. 27202725.

The key ideas underlying the algorithm were discovered by GPT-5.6 Sol Ultra. In fact, HG ran the model for >10 hours each to search for an FPRAS for undirected $s - t$ reliability and a hardness proof for it, and failed at both tasks. WF also had a similar experience independently for undirected $s - t$ reliability. The decisive success came from YF. He ran the model on a prompt adapted from the one OpenAI used for the cycle double cover problem, asking it to remove the acyclicity restriction from the result of Feng and Guo [FG24]. It appears that directly asking for an FPRAS for directed $s - t$ reliability is the right choice, even though it is the harder question. The run lasted 20–21 hours and produced an algorithm based on a Markov chain that samples from a distribution more complicated than the one in Definition 4. That construction is not used verbatim here, but two of its central ideas are retained: the annealing scheme and the flow used to analyse the Markov chain.

For consistency with standard mathematical exposition, the words “we” and “our” are used throughout the paper, including when presenting ideas that originate in the output of the model. The authors verified and simplified the algorithm and the Markov chain proposed by the model, improved the running time, and wrote all of the proofs. They take full responsibility for the correctness of every statement and argument in the paper. GPT-5.6 Sol and Opus 5 were also used to edit the text.

References

- [ACJR21] Marcelo Arenas, Luis Alberto Croquevielle, Rajesh Jayaram, and Cristian Riveros. #NFA admits an FPRAS: Efficient enumeration, counting, and uniform generation for logspace classes. *J. ACM*, 68(6):48:1–48:40, 2021.
- [AvBGM25] Antoine Amarilli, Timothy van Bremen, Octave Gaspard, and Kuldeep S. Meel. Approximating queries on probabilistic graphs. *Log. Methods Comput. Sci.*, 21(4):30:1–30:31, 2025.
- [Bal80] Michael O. Ball. Complexity of network reliability computations. *Networks*, 10(2):153–165, 1980.
- [Bal86] Michael O. Ball. Computational complexity of network reliability analysis: An overview. *IEEE Trans. Reliab.*, 35(3):230–239, 1986.
- [BP83] Michael O. Ball and J. Scott Provan. Calculating bounds on reachability and connectedness in stochastic networks. *Networks*, 13(2):253–278, 1983.

- [CFJ⁺25] Xiaoyu Chen, Weiming Feng, Zhe Ju, Tianshun Miao, Yitong Yin, and Xinyuan Zhang. Faster mixing of the Jerrum-Sinclair chain. In *FOCS*, pages 1007–1028. IEEE, 2025.
- [CGZZ24] Xiaoyu Chen, Heng Guo, Xinyuan Zhang, and Zongrui Zou. Near-linear time samplers for matroid independent sets with applications. In *RANDOM*, pages 32:1–32:12, 2024.
- [Col87] Charles J. Colbourn. *The Combinatorics of Network Reliability*, volume 4 of *The International Series of Monographs on Computer Science*. Oxford University Press, New York, 1987.
- [DGGJ04] Martin E. Dyer, Leslie Ann Goldberg, Catherine S. Greenhill, and Mark Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, 2004.
- [DS91] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.
- [FG24] Weiming Feng and Heng Guo. An FPRAS for two terminal reliability in directed acyclic graphs. In *ICALP*, pages 62:1–62:19, 2024.
- [GGL16] Leslie Ann Goldberg, Rob Gysel, and John Lapinskas. Approximately counting locally-optimal structures. *J. Comput. System Sci.*, 82(6):1144–1160, 2016.
- [GJ19] Heng Guo and Mark Jerrum. A polynomial-time approximation algorithm for all-terminal network reliability. *SIAM J. Comput.*, 48(3):964–978, 2019.
- [GM20] Amir Kafshdar Goharshady and Fatemeh Mohammadi. An efficient algorithm for computing network reliability in small treewidth. *Reliab. Eng. Syst. Saf.*, 193:106665, 2020.
- [Har60] Theodore E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56(1):13–20, 1960.
- [Jer81] Mark Jerrum. *On the Complexity of Evaluating Multivariate Polynomials*. Ph.D. dissertation, Department of Computer Science, University of Edinburgh, 1981. Technical Report CST-11-81.
- [JS89] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989.
- [JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, 2004.
- [Kan94] Ravi Kannan. Markov chains and polynomial time algorithms. In *FOCS*, pages 656–671, 1994.
- [Kar01] David R. Karger. A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem. *SIAM Rev.*, 43(3):499–522, 2001.

- [KL85] Richard M. Karp and Michael Luby. Monte-Carlo algorithms for the planar multiterminal network reliability problem. *J. Complexity*, 1(1):45–64, 1985.
- [LP17] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, Rhode Island, second edition, 2017. With contributions by Elizabeth L. Wilmer.
- [PB83] J. Scott Provan and Michael O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. Comput.*, 12(4):777–788, 1983.
- [Pro86] J. Scott Provan. The complexity of reliability computations in planar and acyclic graphs. *SIAM J. Comput.*, 15(3):694–702, 1986.
- [Sin92] Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.*, 1(4):351–370, 1992.
- [SW85] A. Satyanarayana and R. Kevin Wood. A linear-time algorithm for computing K -terminal reliability in series-parallel networks. *SIAM J. Comput.*, 14(4):818–832, 1985.
- [Val79] Leslie G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comput.*, 8(3):410–421, 1979.
- [ZL11] Rico Zenklusen and Marco Laumanns. High-confidence estimation of small s - t reliabilities in directed acyclic networks. *Networks*, 57(4):376–388, 2011.