Algorithmic, probabilistic, and physics-inspired methods for cuts and shortest paths in graphs

Balázs Ferenc Mezei

Department of Mathematics
Royal Holloway, University of London

This dissertation is submitted for the degree of
Doctor of Philosophy

September 2020

I would like to dedicate this thesis to my loving family ...

Declaration

These doctoral studies were conducted under the supervision of Professor Stefanie Gerke and Professor Gregory Sorkin.

The work presented in this thesis is the result of original research I conducted, in collaboration with others, whilst enrolled in the Department of Mathematics as a candidate for the degree of Doctor of Philosophy. This work has not been submitted for any other degree or award in any other university or educational establishment.

The content of Chapter 2 is joint work with Stefanie Gerke and Gregory Sorkin.

The content of Chapter 3 is joint work with Gregory Sorkin.

The content of Chapter 4 is joint work with Amin Coja-Oglan, Philipp Loick and Gregory Sorkin.

Balázs Ferenc Mezei
September 2020

Acknowledgements

I would like to thank my supervisors, Stefanie Gerke and Gregory Sorkin for always making me feel welcome, their guidance in mathematics and the community thereof, and the many hours spent thinking and drinking coffee together. I have felt incredibly lucky to have them as supervisors, and would also like to thank them for their support outside mathematics.

I am grateful to Amin-Coja Oghlan for the research visits and introducing me to statistical physics.

Thank you to András Sándor, Eszter Rónai, Christina Clarke, Philip Ovington, Nicola Wendt, Eammon Postlethwaite, Jeroen Pijnenburg, Liam Medley, Catherine Keele and Nicholas Winstone. In trying to keep this list short, many friends and teachers were omitted, but are remembered.

Finally, I would like to acknowledge my parents Ágnes and László, and my brother Tamás for their continued support throughout my studies.

Abstract

This thesis considers classical combinatorial optimisation problems in various settings. Optimisation is fundamental in many areas, from computer science, to operations research, biology, chemistry, social sciences and engineering. Broadly speaking, combinatorial optimisation problems consist of finding an optimal solution or configuration from a large (but finite) set of possibilities.

Chapter 2 studies the shortest path problem in a random setting. Consider a complete graph K_n with edge weights drawn independently from a uniform distribution $U(0, 1)$. The weight of the shortest (minimum-weight) path P_1 between two given vertices is known to be $\ln n/n$, asymptotically almost surely almost exactly (a.a.s. a.e.). We define a second-shortest path P_2 to be the shortest path edge-disjoint from P_1, and consider more generally the shortest path P_k edge-disjoint from all earlier paths. We show that the cost X_k of P_k is a.a.s. a.e. $(2k + \ln n)/n$, uniformly for all $k \leq n - 1$. We show analogous results when the edge weights are drawn from an exponential distribution.

Chapters 3 and 4 study partitioning problems on d-regular graphs. In Chapter 3, we show improved bounds on the minimum bisection size (bisection width) for *arbitrary* d-regular graphs by analysing a local iterative algorithm, obtaining improved asymptotic bounds for $5 \leq d \leq 125$. In Chapter 4, we show improved upper bounds on the maximum cut size of *random* d-regular graphs by analysing the Ising antiferromagnet, using mathematically rigorous techniques drawn from statistical physics.

Table of contents

Chapter 1

Introduction

A graph is a simple mathematical structure that is frequently used by mathematicians, computer scientists, engineers, social scientist, biologists, chemists and others to model problems. A wide range of complex real-world problems and models can be captured accurately by representing them as graphs, e.g. social networks, electrical circuit layouts, load balancing, transportation networks, spread of diseases, logistics problems. Especially since the advent of computing, the study of the structures of (random) graphs, and graph algorithms have been under constant research.

Formally, a (simple) graph $G = (V, E)$ is a set of vertices V and edges $E \subseteq V^{(2)}$, where $V^{(2)}$ denotes the set of all unordered pairs of V. Sometimes, each edge $e \in E$ is assigned a weight, in which case we call G a weighted graph. For example, the weight can represent the distance between two cities, or the strength of a connection in a network.

This thesis contributes to the field of combinatorial optimization in three different settings. We give a brief introduction to each in the following subsections.

1.1 Successive shortest paths in K_n with random edge weights

Dijkstra's algorithm is one of first graph algorithms, published in 1959 [Dij59]. Given a graph G with non-negative edge weights, it finds the shortest path between two vertices of G. We study the shortest path problem in a random setting.

Let $G = K_n$ be the complete graph on n vertices, and assign i.i.d. uniform $(0, 1)$ weights to each edge. Fixing vertices $s, t \in V(G)$ it is natural to consider the weight of the cheapest path P_1 between s and t. It is a well-known result of Janson [Jan99] that X_1, the cost of P_1, is asymptotically almost surely almost exactly (a.a.s. a.e.) $\log n/n$. That is, as $n \to \infty$,

$$\frac{X_1}{\ln n/n} \xrightarrow{\text{p}} 1. \tag{1.1}$$

We define the second cheapest path P_2, with cost X_2, to be the cheapest s–t path edge-disjoint from P_1, and in general define P_k, with cost X_k, to be the cheapest s–t path edge-disjoint from $P_1 \cup \cdots \cup P_{k-1}$, provided such a path exists (for $k \leq n/2$, the path P_k is deterministically guaranteed to exist, but for $k > n/2$, there is a small probability that it does not).

Our question is how the costs X_k behave in the limit as $n \to \infty$. Note that P_k can only exist for $k \leq n-1$, as s and t have degree $n-1$. We generalise the result of eq. (1.1) all the way to $k \leq n-1$: we show that for $k \leq n-1$, the cost of P_k is a.a.s. a.e. $2k/n + \log n/n$. More precisely, we show that

$$\frac{X_k}{2k/n + \ln n/n} \xrightarrow{\text{p}} 1 \tag{1.2}$$

uniformly in k.

For similar problems, it is common that the results automatically extend to other edge weight distributions, however this is not the case here. Nonetheless, we give an extension to the case where the edge weights are i.i.d. exponentially distributed. For both distributions, we also establish the expectation of X_k conditioned on the existence of P_k (as it may not exist for $k > n/2$).

Similar questions on the minimum cost of other structures have been studied: minimum spanning trees [Fri85], matchings in bipartite graphs [Ald01; Wäs09], Hamilton cycles [Fri04], and travelling salesmen tours [Wäs10]. Recently, questions on successive and simultaneous structures have been explored for minimum spanning trees [FJ18; JS19], motivating our work.

1.2 Partitioning problems

Partitioning the graph into smaller pieces under certain constraints is a fundamental algorithmic and theoretical challenge. The study of partitioning problems has given rise to a rich literature and contribution to complexity theory. It has also lead to the development of important algorithmic techniques, such as network flows and semidefinite programming [ARV04; GW95]. Practical uses include divide and conquer algorithms [Shm97; LLR95], cost minimisation in VLSI design [AK95], and parallel computing [HL95]; for a survey see [PT93].

We study a special case of graph partitioning problems, specifically we study the minimum bisection and maximum cut of d-regular graphs. A *cut* of a graph G is a partition of the vertex set into two parts, and the *size* of a cut is the number of edges between the two parts. A bisection is a cut with the two parts as equal as possible in size. The minimum bisection of a graph G, denoted $\mathrm{MinBis}(G)$, is the size of the bisection with minimum cut size; the maximum cut, denoted $\mathrm{MaxCut}(G)$, is the size of the cut with maximum cut size. The minimum bisection is also known as the bisection width, as we will refer to it in Chapter 3.

Finding the minimum bisection and maximum cut of a graph are both NP-hard problems, therefore unless $\mathrm{P} = \mathrm{NP}$, there is no polynomial time algorithm to find them. They remain NP-hard, even when restricted to d-regular graphs, for $d \geq 3$ [GJ09; BCLS87; GJS76].

1.2.1 Minimum bisection (bisection width)

As mentioned before, the partitioning of a system to minimise the connection between the two clusters (parts) arises naturally in VLSI design [AK95]. Current chip designs now have transistor counts in excess of billions, and partitioning these systems to smaller components makes them more manageable for both design and production. Interactions still need to happen between the components, and generally one wishes to minimise these as they increase manufacturing costs and have a negative impact on performance. Consider the graph with the n transistors as vertices, and an edge between two transistors if interaction between them is required. The optimal allocation of transistors (the one that minimises interactions) corresponds exactly to the minimum bisection of the this graph.

Even approximating the minimum bisection is difficult: the best known algorithmic guarantee in polynomial time is an approximation within a factor of $\mathcal{O}(\log n)$ [Räc08].

In Chapter 3 we derive upper bounds on the minimum bisection size for arbitrary d-regular graphs. We do this by analysing an iterative local search algorithm: we start with any arbitrary bisection of G, then iteratively move a small set of vertices from the current larger part to the smaller, decreasing the cut size in each step. When we can no longer find such a small set, we naively rebalance the partition to obtain a bisection.

The method is based on an earlier paper by Monien and Preis [MP06] which deals with the $d = 3, 4$ cases. We generalise the technique for all d, and we obtain improved asymptotic upper bounds on the bisection width of arbitrary d-regular graphs with $5 \leq d \leq 125$.

1.2.2 Maximum cut

We study the maximum cut size of a random d-regular graph $G(n,d)$, which is a random graph sampled uniformly from all d-regular graphs on n vertices. (Note that this only exists when dn is even.) Specifically, we derive high probability upper bounds on the random variable $\mathrm{MaxCut}(G(n,d))$, by analysing the *Ising antiferromagnet*, described below.

Given a graph $G=(V,E)$, a mapping $\sigma \in \{\pm 1\}^V$ of each vertex to a spin $+1$ or -1 is called a *configuration*. A configuration σ naturally gives rise to a cut of G, by letting the two vertex classes be the vertices to which σ assigns $+1$ and -1 respectively. Define the *total energy* (also known as the *Hamiltonian*) of a configuration σ as

$$\mathcal{H}_G(\sigma) = \sum_{uv \in E(G)} \mathbf{1}\{\sigma_u = \sigma_v\},$$

so that $\mathcal{H}_G(\sigma)$ counts the number of uncut (unsatisfied) edges in the cut given by σ.

The *Ising antiferromagnet* with inverse temperature (parameter) $\beta \geq 0$ on G is a probability distribution $\mu_{G,\beta}$ on all configurations $\Omega = \{\pm 1\}^V$ defined by

$$\mu_{G,\beta}(\sigma) = \frac{1}{Z_\beta(G)} \exp(-\beta \mathcal{H}_G(\sigma)),$$
$$Z_\beta(G) = \sum_{\sigma \in \{\pm 1\}^V} \exp(-\beta \mathcal{H}_G(\sigma)).$$

The normalising constant $Z_\beta(G)$ is called the *partition function*. The distribution assigns a penalty term $\exp(-\beta)$ for each unsatisfied (uncut) edge in the configuration σ. At $\beta = 0$, the distribution becomes the uniform distribution on all configurations. In the limit $\beta \to \infty$, the probability becomes concentrated on configurations with minimal energy, i.e. those corresponding to the maximum cut of G. This is the limit we will be interested in, as we have

$$\mathrm{MaxCut}(G) = \frac{dn}{2} - \min_{\sigma \in \{\pm 1\}^V} \mathcal{H}_G(\sigma) \leq \frac{dn}{2} + \frac{1}{\beta} \log Z_\beta(G).$$

We obtain a (high probability) upper bound on $\mathrm{MaxCut}(G(n,d))$ from the last expression as follows. We show an upper bound on the expectation of the log-partition function, $\mathbb{E}[\log Z_\beta(G(n,d))]$ (which in this case will be a negative quantity), and use that $\log Z_\beta(G(n,d))$ is tightly concentrated around its mean.

The technique we apply is called the *interpolation method*. It has recently become standard in literature, and has been applied to numerous problems in combinatorics and theoretical computer science: to establish lower bound on the chromatic number of random graphs [ACG19], to establish the satisfiability threshold for random k-SAT [DSS15], to count the number of solutions to random k-NAE-SAT [SSZ16] and to establish an upper bound independence number of random graphs [LO18].

Chapter 2

Successive shortest paths in K_n with random edge weights

2.1 Introduction

It is a standard problem to find the shortest s–t path in a graph, i.e., the cheapest path P_1 between specified vertices s and t, and its cost X_1, where the cost of a path is the sum of the costs of its edges. We will use the terms “cost” and “weight” interchangeably, and reserve “length” for the number of edges in a path. Shortest paths will refer to paths of minimum weight, however we will use “short” to describe a path of small length.

Consider the complete graph $G = K_n$ with each edge $\{u, v\}$ having weight $w(u, v)$, where the $w(u, v)$ are i.i.d. random variables with exponential distribution Exp(1) or uniform distribution $U(0, 1)$ (we consider both versions). In this random setting, a well-known result of Janson [Jan99] is that as $n \to \infty$,

$$\frac{X_1}{\ln n/n} \xrightarrow{\text{p}} 1. \tag{2.1}$$

We define the second cheapest path P_2, with cost X_2, to be the cheapest s–t path edge-disjoint from P_1, and in general define P_k, with cost X_k, to be the cheapest s–t path edge-disjoint from $P_1 \cup \cdots \cup P_{k-1}$, provided such a path exists. We also think of this as finding path P_k after the preceding paths’ edges have been removed. Our question is how the costs X_k behave in the limit as $n \to \infty$ (this limit is implicit throughout). Our main result is the following.

Theorem 2.1.1. *In the complete graph K_n with i.i.d. uniform $U(0, 1)$ edge weights, with X_k the cost of the kth cheapest path,*

$$\frac{X_k}{2k/n + \ln n/n} \xrightarrow{\mathrm{p}} 1 \tag{2.2}$$

uniformly for all $k \leq n-1$. That is, for any $\varepsilon > 0$, asymptotically almost surely, for every $k = 1, \dots, n-1$,

$$1 - \varepsilon \leq \frac{X_k}{2k/n + \ln n/n} \leq 1 + \varepsilon. \tag{2.3}$$

Naturally, with $k = 1$, eq. (2.2) recovers Janson's result eq. (2.1), since $2/n = o(\ln n/n)$.

As discussed shortly, in contrast to many cases, the result for the uniform distribution does not extend immediately to all distributions with positive density at 0. However, we have a corresponding result for exponentially distributed edge weights. Given an edge-weight distribution, let $W_{(k)}$ be the (random) weight of the kth cheapest edge out of a vertex (the kth order statistic of $n-1$ edge weights).

Theorem 2.1.2. *In the complete graph K_n with i.i.d. exponential edge weights with mean 1,*

$$\frac{X_k}{2\,\mathbb{E}\,W_{(k)} + \ln n/n} \xrightarrow{\mathrm{p}} 1 \tag{2.4}$$

uniformly for all $k \leq n-1$.

We give the guiding intuition behind the formula eq. (2.4) in Section 2.1.1. Note that $\mathbb{E}\,W_{(k)} = \sum_{i=1}^{k} \frac{1}{n-i}$ in the exponential case (see e.g. Lemma 2.4.2). In the uniform case, $\mathbb{E}\,W_{(k)} = k/n$, so eq. (2.2) in Theorem 2.1.1 can also be written as eq. (2.4).

Rather than finding the k successive cheapest paths, we may alternatively wish to find the k edge-disjoint paths of *collective* minimum cost. Equivalently, where every edge of G has capacity 1, we may be interested in the minimum-cost k-flow from s to t in G. The following remark shows that this problem leads to essentially the same costs. (The analogous "collective" problem for minimum spanning trees is solved in [FJ18], and [JS19] shows that for MSTs, the "successive" version leads to strictly larger costs.)

Remark 2.1.3. *In the complete graph K_n with i.i.d. edge weights with distribution $U(0, 1)$ or exponential with mean 1, the minimum-cost k-flow has cost F_k satisfying*

$$\frac{F_k}{\sum_{i=1}^{k}(2\,\mathbb{E}\,W_{(i)} + \ln n/n)} \xrightarrow{\mathrm{p}} 1 \tag{2.5}$$

uniformly for all $k \leq n-1$.

As in eq. (2.3), the statement consists of high-probability upper and lower bounds. The upper bounds here, for the two models, follow immediately from the upper bounds of eq. (2.2) and eq. (2.4). The lower bounds follow from the lower bound on $S_k := \sum_{i=1}^{k} X_k$ (see eq. (2.84)) in eq. (2.87) and its analogue for the exponential case, as those bounds hold for any set of k edge-disjoint paths. (The main work in Section 2.7, not needed here, is to extract lower bounds on X_k from the lower bounds on S_k.)

Remark 2.1.4. *P_k is always defined for all $k \leq n/2$, but, at least for n even, may be undefined for all $k > n/2$.*

Proof. There are $n-2$ length-2 s–t paths. Any path P_k can destroy (share an edge with) at most two such paths (since P_k uses just one edge incident to each of s and t). Also, the single-edge path $\{s,t\}$ is destroyed only by the path P_k consisting of just this edge. So, for $P_1, \dots, P_k$ to destroy all length-1 and length-2 paths requires $k \geq (n-2)/2 + 1 = n/2$, so for $k \leq n/2$, certainly path P_k exists.

Conversely, a construction described in 1892 by Lucas [Luc92, pp. 162–164], which he attributes to Walecki, shows that a complete graph K_{2r} can be decomposed into r edge-disjoint Hamilton paths (whose $2r$ terminals are all distinct). For n even, decompose $G = K_n \setminus \{s,t\}$ in this way, then link s to one "start" terminal of each such path and t to the other "end" terminal, giving $(n-2)/2$ edge-disjoint s–t paths. The edge $\{s,t\}$ gives another path, for $n/2$ paths in all. The only edges not used by these paths are a star from s to the Hamilton paths' end terminals, and another star from t to their start terminals, and as there are no other unused edges to connect these two stars, there is no further s–t path. With nonzero probability, the edge weights are such that $P_1, \dots, P_{n/2}$ are these $n/2$ paths, so that $P_{n/2+1}$ does not exist. □

Remark 2.1.4 implies that, at least for n even, $\mathbb{E}[X_k]$ is undefined for $k > n/2$. The following theorem establishes $\mathbb{E}\, X_k$ for $k \leq n/2$, and for all $k \leq n-1$, gives the expectation conditioned on the (high-probability) event that P_k exists.

Theorem 2.1.5. *In both the uniform and exponential models, for $k \leq n-1$, a.a.s. P_k exists, and*

$$\mathbb{E}[X_k \mid P_k \text{ exists}] = (1+o(1))(2\, \mathbb{E}\, W_{(k)} + \ln n/n), \tag{2.6}$$

uniformly in k.

For $k \leq n/2$, by Remark 2.1.4 the conditioning is null, so it is immediate from Theorem 2.1.5 that $E[X_k] = (1+o(1))(2\, \mathbb{E}\, W_{(k)} + \ln n/n)$.

2.1.1 Intuition

The intuitive picture is that path P_k should use the kth cheapest edges out of s and t, whose costs are denoted $W^s_{(k)}$ and $W^t_{(k)}$ respectively. Then, if we ignore previous paths' use of other edges in $G \setminus \{s,t\}$, by eq. (2.1) the opposite endpoints of these two edges should be connected by a path of cost about $\ln n/n$. This suggests that $X_k \leq W^s_{(k)} + W^t_{(k)} + \ln n/n$, and this is our guiding intuition. Obviously, the path P_k does not have to use the kth cheapest edge, its middle section may cost more or less than $\ln n/n$, and as earlier paths use up edges, the costs of these middle sections may rise. It is true, though, that $\sum_{i=1}^k X_i \geq \sum_{i=1}^{k-1} \left(W^s_{(k)} + W^t_{(k)}\right)$ (summing only to $k-1$ on the right-hand side to avoid doubly counting edge $\{s,t\}$), and we use this in proving the lower bounds on X_k (in Section 2.7 for uniform and Section 2.8.7 for exponential) and, more surprisingly, in proving the upper bounds on X_k for large k (in Section 2.5 generically, the details treated in Sections 2.6 and 2.8).

Our upper bounds are obtained by reasoning as follows. Janson [Jan99] analyses the shortest s–t path, and shortest-path tree (SP tree or SPT) on s, in the randomly edge-weighted graph $G = K_n$, showing that the cost of P_1 is asymptotically almost surely, almost exactly $\ln n/n$. When the path P_1 is deleted, this prunes away a root-level branch of the SP tree. The SP tree is a uniform random tree, and using known properties of such trees (see for example [SFH06]) it is not hard to show that what remains of the SP tree is likely to be large; capitalising on this we can find an almost equally cheap path P_2'. This line of argument also shows that there remains a cheap path after deleting P_2', but we need to know what happens when we delete the true second-shortest path P_2, and at this point the argument fails because it gives no characterisation of P_2, only of P_2'. We do know, however, that P_2 is cheap (no more expensive than P_2'), and of course uses just one edge incident to each of s and t, and we will show that deleting *any* edge set with these properties (including P_2 as a possibility) must still leave a cheap path P_3', and so forth. This "adversarial" deletion argument is developed in Section 2.3.2 to prove Theorem 2.1.1.

2.1.2 Context

The question fits with a broad research theme on optimisation (and satisfiability) problems on random structures. The novel element here is the "robustness" aspect of finding cheap structures even after the cheapest has been removed, and in this we were motivated by a recent study by Janson and Sorkin [JS19] of the same question for successive minimum spanning trees (MSTs), again for K_n with uniform or exponential random edge weights. The results for shortest paths and MSTs are dramatically different. For MSTs, it is a celebrated result of Frieze [Fri85] that as $n \to \infty$ the cost of the MST T_1 satisfies $w(T_1) \xrightarrow{\text{p}} \zeta(3) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} 1/k^3$, and [JS19] shows that each subsequent tree's cost has $w(T_k) \xrightarrow{\text{p}} \gamma_k$ with the γ_k strictly increasing (and $2k - 2\sqrt{k} < \gamma_k < 2k + 2\sqrt{k}$). That is very different from the case here, for paths, where for $k = o(\ln n)$ we have X_k asymptotically equal to X_1.

Further context is given in the discussion of open problems in Section 2.2.2.

2.1.3 Edge weight distributions

As remarked earlier, in many contexts (including for the length X_1 of a shortest path) the result for any distribution with positive density at 0 follows immediately from that for the uniform distribution $U(0, 1)$, but that is not the case for the successive paths considered here.

Remark 2.1.6. *Janson proves the X_1 case in the exponential model but provides standard "black-box" reasoning that it holds also for the uniform distribution, for any distribution with density 1 at 0 (i.e., with cumulative distribution function $F(x) = \mathbb{P}(X \leq x) = x + o(x)$ for $x \searrow 0$), and, after simple rescaling, for any distribution with positive density at 0. Simply, if there is a path P of cost $w(P) = o(1)$ in some such model, each edge $w \in P$ must have cost $w(e) \leq w(P)$. Coupling with the uniform distribution by replacing w with $w' = F(w)$, for $w \leq w(P) = o(1)$ we have that*

$$w' = (1 + o(1))w,$$

and thus the same path P is similarly cheap in the uniform model:

$$w'(P) = \sum_{e \in P} w'(e) = (1 + o(1)) \sum_{e \in P} w(e) = (1 + o(1))w(P). \tag{2.7}$$

By the same token, if a path is cheap in any model, the same path has asymptotically the same cost in any other model, and thus the cheapest paths have asymptotically the same cost.

Remark 2.1.7. *In our setting this argument does not apply: to find path P_k we must know the nature of the $k - 1$ previous paths; their costs are not enough. For $k = o(n)$, however, the standard argument applies within our proofs, since the proofs rely only on edges of cost $o(1)$. However, for larger k there are genuine difficulties. Our argument for the exponential case, in Section 2.8, largely parallels that for uniform but requires new calculations for the upper bound, and one new idea for the lower bound (in Section 2.8.7). It is not clear for what other edge-weight distributions (even those with density 1 at 0) eq. (2.4) will hold.*

2.2 Open problems

2.2.1 Poisson multigraph model

The issue of possible non-existence of paths P_k for $k > n/2$ (see Remark 2.1.4) is obviated if, as in [JS19], we work in a Poisson multigraph model. Here, each pair of vertices $\{u, v\}$ of K_n is joined by infinitely many edges, whose weights are drawn from a Poisson process of rate 1 (so that the cheapest $\{u, v\}$ edge has exponentially distributed cost of mean 1). By construction, in this model every s–t path is always available (possibly at a higher cost).

Conjecture 2.2.1. *In the Poisson multigraph model,* $\frac{X_k}{2k/n+\ln n/n} \xrightarrow{\text{p}} 1$ *uniformly for all* $k \leq n-1$*, and* $\frac{\mathbb{E} X_k}{2k/n+\ln n/n} \to 1$ *for all* $k \leq n-1$.

Actually, in this model there is no need to stop at $k = n-1$, but it is not clear how far out we can go (especially preserving uniform convergence).

2.2.2 Other models

Most narrowly, it would be interesting to characterise successive shortest paths that are vertex-disjoint rather than edge-disjoint, and (in the style of Remark 2.1.3 for edge-disjoint paths) the k vertex-disjoint paths of collective minimum cost. In this model, guessing that path lengths stay around $\log n$, we would expect P_k to be defined up to k about $n/\log n$.

More broadly, it would be interesting to explore different edge-weight distributions, different structures, and different graphs.

As noted earlier, we have results for uniformly and exponentially distributed edge weights, but not for arbitrary distributions. As mentioned, results for the single shortest path follow by standard arguments for any distribution with positive density near 0. For a distribution with density tending to 0 or ∞ at 0, shortest paths were studied in [BH12]. In particular, they consider the case when edge weights are i.i.d. and have the same distribution as Z^p, where $Z \sim \text{Exp}(1)$ and $p > 0$ is a fixed parameter; in this setting, the shortest path has length $p \ln n$ and its cost is $\ln n/n^p$ times a p-dependent constant. A variant where the edge-weight distribution may depend on n is studied in [EGvdHN13].

To what distributions does Theorem 2.1.2 extend? Restricting to distributions with positive density near 0, the arguments in Section 2.8 should immediately extend for all $k = o(n)$. For larger k, the "middle" of each path should remain short, so the issue is the edges incident on s and t in P_k. Certainly eq. (2.4) will fail if the order statistics of edges incident to s are not concentrated, for example if the edge distribution is a mixture of $U(0, 1)$ and an atom at 2 or (for a continuous example) a mixture of $U(0, 1)$ and the

Pareto distribution with CDF $1 - 1/x$ for $x \geq 1$. It might be true that eq. (2.4) holds more generally if the expectation $2\,\mathbb{E}\, W_{(k)}$ is replaced by $W^s_{(k)} + W^t_{(k)}$. However, to obtain the needed lower bound for the exponential model (see Section 2.8), we had to address the the fact that the kth path does not necessarily use the edges of cost $W^s_{(k)}$ and $W^t_{(k)}$; we also needed exponential-specific calculations for the upper bound.

One could explore other structural models. Minimum spanning trees (MSTs) have already been explored in [JS19] for the successive version and in [FJ18] for the collective version. But for many other models the single cheapest structure is well studied but the successive and collective extensions have not been explored: this includes perfect matchings in complete bipartite graphs $K_{n,n}$ [Ald01; Wäs09], perfect matchings in complete graphs K_n [Wäs08], and Hamilton cycles (i.e., the Travelling Salesman Problem) in K_n [Wäs10].

One could also consider graphs other than complete graphs, in the style of studies of the MST in a random regular graph [BFM98], and of first-passage percolation in Erdős–Rényi random graphs [BHH11] and hypercubes [Mar16].

2.3 Upper bound for small k

In this section we prove the upper bound of Theorem 2.1.1 for all $k = o(\sqrt{n})$; larger values are treated in the next section.

As discussed in the introduction, we can characterise the cheapest path P_1 and subsequent paths that are *cheap* but not necessarily *cheapest*, putting us at a loss to characterise what remains on deletion of a subsequent *cheapest* path. We address this in this section. Given k, we show a construction of a subgraph $R = R^{(k)}$ of G designed so that, as we will show in turn, its s–t paths are all cheap, and no deletion of edges from R subject to certain constraints can destroy all these paths. We show that the union of the k shortest paths satisfies these constraints, so that there remains a cheap s–t path in R and thus in G, and use this to prove Theorem 2.1.1.

Specifically, we will define a structure R, sketched in Fig. 2.1, that has many cheap and spread-out paths between s and t, within which we will always find a cheap path. A crucial point is that each step of the construction occurs in a complete induced subgraph of G of size $n - o(n)$ with all edges unconditioned.

We will show, assuming that

$$X_i \leq (1+\varepsilon)\left(\frac{2i}{n} + \frac{\ln n}{n}\right) \tag{2.8}$$

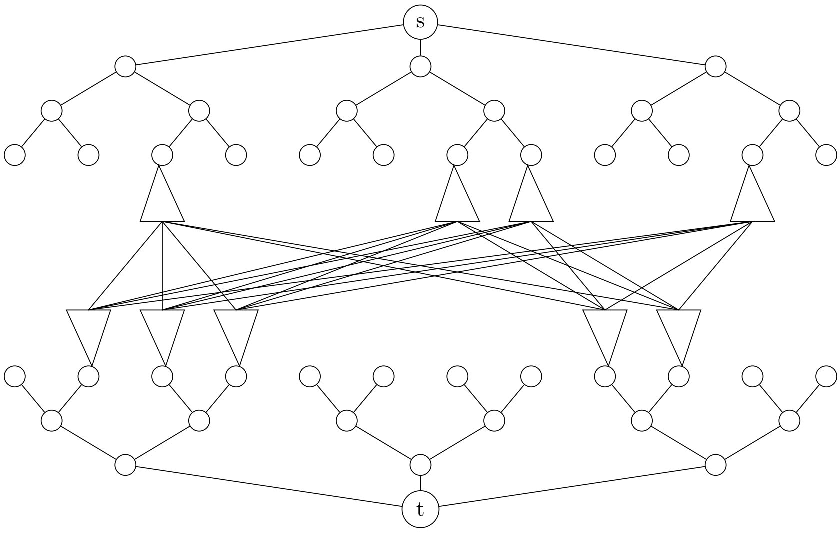


Fig. 2.1 Cartoon of a robust subgraph R of G, showing the vertices s and t, their respective structures R_s and R_t including shortest-path trees represented by triangles (some "failed" and thus not shown), and the cheap edges connecting triangles in R_s and R_t. Vertices s and t have down-degree (number of children) r_0, and vertices at levels 1 and 2 (in R_s and R_t) have down-degrees r_1 and r_2 respectively.

for all $i \leq k$, that the same holds for $i = k+1$. We will do so by showing that after deleting k paths, each of cost $\leq (1+\varepsilon)(2k/n + \ln n/n)$ from G, some or all of whose edges may lie in R, there remains a path in R satisfying the same cost bound, and so this must also be true of P_{k+1}.

Consistent with this approach, and because to prove convergence in probability it suffices to consider an arbitrarily small, fixed ε (see around eq. (2.3)), throughout this section we assume that $\varepsilon > 0$ is fixed. Thus, in the $n \to \infty$ limit implicit throughout,

$$\varepsilon = \Theta(1), \tag{2.9}$$

and ε (and functions of ε) may be absorbed into the constants implicit in any Landau-notation expression.

Remark 2.3.1. *Most of the calculations below hold for any $\varepsilon > 0$, but a few (eq. (2.14) and eq. (2.15) for example) hold only for ε sufficiently small. This is not restrictive here, in proving convergence in probability, but to characterise expectation, Section 2.9.1 requires ε to be a large constant (to assure sufficiently small failure probabilities). The proof of Lemma 2.9.1 addresses the changes needed.*

Before going into detail let us sketch the construction of R. We first build up a tree R_s on s, starting from s at level 0, the opposite endpoints of edges out of level i forming level $i+1$. We will always choose "cheap" edges, but not always the cheapest ones, as explained later. From s we will choose $k+r_0$ cheap edges; from each of these $k+r_0$ level-1 vertices we choose r_1 cheap edges; from each of the $(k+r_0)r_1$ level-2 vertices we choose r_2 cheap edges; and on each of the $(k+r_0)r_1r_2$ level-3 vertices we construct a shortest-path tree comprising d vertices. We do a similar construction on t to form R_t. Finally, we link R_s and R_t using cheap edges between their shortest-path trees. The values of the parameters r_0, r_1, r_2 and d are given in eq. (2.18), eq. (2.20), eq. (2.24) and eq. (2.26), and it is confirmed in Section 2.3.9 that the construction uses only a small fraction of G's vertices,

$$|V(R)| = \mathcal{O}((k+r_0)r_1r_2d) = o(n), \tag{2.10}$$

a fact we rely on in the construction.

We will repeatedly use the following Chernoff bound, which in fact holds under more general conditions; see for example [Jan02, Theorem 1, eq. (4)].

Lemma 2.3.2. *Let $X \sim \mathrm{Bi}(n,p)$ be a binomial random variable with mean $\lambda = np$. Then for any $\varepsilon > 0$, $\mathbb{P}(X < (1-\varepsilon)\lambda) \le \exp(-\varepsilon^2\lambda/2)$.*

2.3.1 Cheap paths are short

We show that, w.h.p., every cheap path in G is also short. The following lemma asserts the contrapositive. The result is used in eq. (2.16) to restrict the number of edges the adversary can delete.

Lemma 2.3.3. *In both the uniform and exponential models, with probability $1-\mathcal{O}(n^{-1.9})$, simultaneously for all l with $\ln n \le l < n$, every s–t path of length l has cost $\ge l/(19n)$.*

Proof. We start with the uniform distribution. Here, with $X = \sum_{i=1}^{l} X_i$, $X_i \sim U(0,1)$ i.i.d., X has the Irwin-Hall distribution and it is a standard result that $\mathbb{P}(X \le a) \le a^l/l!$ (see for example [FPS18, eq. 8]). Recall that Stirling's approximation is also a lower bound. Thus,

$$\mathbb{P}\left(X \le \frac{l}{19n}\right) \le \frac{(l/19n)^l}{l!} \le \frac{(l/19n)^l}{\sqrt{2\pi l}\,(l/e)^l} < \left(\frac{e}{19n}\right)^l .$$

The cost of a fixed path of length l has the same law as X. Over the $\le n^l$ choices for such a path, the number M_l of "cheap paths" (of cost $< l/(19n)$) satisfies (by Markov's inequality)

$$\mathbb{P}(M_l > 0) \le \mathbb{E}\,M_l \le n^l\,\mathbb{P}\left(X \le \frac{l}{19n}\right) \le n^l \left(\frac{e}{19n}\right)^l = \left(\frac{e}{19}\right)^l .$$

Summing over $l \geq \ln n$, the probability that there is a cheap path of any such length is $\mathcal{O}\Big((e/19)^{\ln n}\Big) = \mathcal{O}\big(n^{-1.9}\big)$.

Since an Exp(1) random weight X' can be obtained from a $U(0,1)$ r.v. X by setting $X' = -\ln(1-X) > X$, the exponential weight stochastically dominates the uniform, so the result for uniform immediately implies that for exponential. □

2.3.2 Adversarial edge deletions

As noted in the introduction, we introduce an edge-deleting adversary whose powers allow it to delete the paths $P_1, \dots, P_k$, but which is more easily characterised than those paths are. We now specify what the adversary is permitted to do.

Let

$$s = s(k) \coloneqq 2k + \ln n. \tag{2.11}$$

(From context it should be easy to distinguish this use of s from that as the source of an s–t path.) Let w_0 be the "target cost" of a path, namely

$$w_0 = w_0(k) \coloneqq \frac{s}{n} = \frac{2k}{n} + \frac{\ln n}{n}. \tag{2.12}$$

Define a "heavy" edge to be one of cost

$$\geq \tfrac{1}{11}\varepsilon w_0. \tag{2.13}$$

Assuming that each of $P_1, \dots, P_k$ has weight $\leq (1+\varepsilon)w_0$, the *number of heavy edges* in $P_1 \cup \dots \cup P_k$ is at most

$$\frac{k(1+\varepsilon)w_0}{\frac{1}{11}\varepsilon w_0} < \frac{12k}{\varepsilon} < \frac{12s}{\varepsilon}. \tag{2.14}$$

Also, modulo the one-time failure probability $\mathcal{O}\big(n^{-1.9}\big)$ from Lemma 2.3.3, by that lemma each path has length at most

$$(1+\varepsilon)w_0 \cdot 19n < 20s. \tag{2.15}$$

Thus, the length of all k paths taken together (i.e., the number of edges in $P_1 \cup \cdots \cup P_k$) is at most

$$20ks < 10s^2. \tag{2.16}$$

And of course the k paths include

$$\text{exactly } k \text{ edges incident on each of } s \text{ and } t. \tag{2.17}$$

Subject to these assumptions — that each of $P_1, \dots, P_k$ has weight $\leq (1+\varepsilon)w_0$ and that the high-probability conclusion of Lemma 2.3.3 holds — $P_1 \cup \cdots \cup P_k$ satisfies all three of the constraints eq. (2.14), eq. (2.16), and eq. (2.17) on heavy edges, all edges, and "incident" edges. An adversary who can delete any edge set subject to these constraints is able to delete $P_1 \cup \cdots \cup P_k$, which is all we require. However, to simplify analysis we will give the adversary even more power.

At the root of R we will allow the adversary to delete edges subject only to eq. (2.17); at level 1, additional edges subject only to the "heavy-edge budget" eq. (2.14); and at levels 2 and 3 and for "middle" edges, additional edges subject only to the "edge-count budget" eq. (2.16).

We will show how to choose the parameters of R so that every s–t path in R has cost $\leq (1+\varepsilon)w_0$, and so that R is "robust": after the adversarial deletions, at least one path remains. Specifically, we will arrange that there remains a path in which the "root" edge incident to s costs $\leq \frac{k}{n} + \frac{1}{9}\varepsilon w_0$, the edge out of level 1 is heavy but has cost $\leq \frac{1}{9}\varepsilon w_0$, the edge out of level 2 may be light or heavy and also has cost $\leq \frac{1}{9}\varepsilon w_0$, the path through the SP tree has total cost $\leq \frac{1}{2}\frac{\ln n}{n} + \frac{1}{9}\varepsilon w_0$, the central edge joining this to the opposite SP tree adds cost $\leq \frac{1}{9}\varepsilon w_0$, and the continuation of this path to t has the symmetrical properties. It is immediate that such a path has total cost $\leq (2k + \ln n)/n + 9 \cdot \frac{1}{9}\varepsilon w_0 = (1+\varepsilon)w_0$. (But see eq. (2.35) for confirmation, after the construction is detailed.)

Note: We use the variable n' in the following sections to denote the number of vertices not currently in R. This changes after each level is built, but the crucial point is that we always have $n' = (1 - o(1))n$, which we confirm in Section 2.3.9.

2.3.3 Level 0, cheapest edges

On s, add to R the $k + r_0$ edges of lowest cost, excluding $\{s, t\}$ from consideration, with

$$r_0 = \left\lceil \frac{1}{10}\varepsilon s \right\rceil = \Theta(s). \tag{2.18}$$

Consider this step a *failure* if any selected edge has cost greater than $\frac{k}{n} + \frac{1}{9}\varepsilon w_0$. There are $n' = n - 2 = (1 - o(1))n$ edges under consideration, with weights i.i.d. $U(0, 1)$, and failure occurs iff the number X of edges with weights in the interval $[0, \frac{k}{n} + \frac{1}{9}\varepsilon w_0]$ is smaller than $k + r_0$. Note that $X \sim \text{Bi}(n', \frac{k}{n} + \frac{1}{9}\varepsilon w_0)$, thus $\mathbb{E} X = (1 - o(1))\,(k + \frac{1}{9}\varepsilon s)$, and failure means that $X < k + r_0$, i.e., that

$$\frac{X}{\mathbb{E} X} < (1 + o(1))\,\frac{k + r_0}{k + \frac{1}{9}\varepsilon s} = (1 + o(1))\,\frac{k + \frac{1}{10}\varepsilon s}{k + \frac{1}{9}\varepsilon s},$$

which by $s > 2k$ is

$$< (1 + o(1))\,\frac{k + \frac{1}{10}\varepsilon \cdot 2k}{k + \frac{1}{9}\varepsilon \cdot 2k} = (1 + o(1))\,\frac{(1 + \frac{2}{10}\varepsilon)k}{(1 + \frac{2}{9}\varepsilon)k} < 1 - \frac{1}{50}\varepsilon = 1 - \Omega(\varepsilon).$$

By Lemma 2.3.2, then, the probability of failure is

$$\mathbb{P}(X < (1 - \Omega(\varepsilon))\,\mathbb{E} X) \le \exp(-\Omega(\varepsilon^2)\,\mathbb{E} X/2) \le \exp(-\Omega(\varepsilon^2 \cdot \varepsilon s)) \le \exp(-\Theta(s)), \tag{2.19}$$

the final expression using that ε is constant (see (2.9)).

So, modulo the given failure probability, every selected edge incident on s has cost $\le \frac{k}{n} + \frac{1}{9}\varepsilon w_0$, and after the adversarial deletion of k of these edges, r_0 remain. The selection of edges conditions the costs of the other edges incident on s, but none will play any role in the analysis.

The purpose of the next two levels is to expand the number of edges to the point where the adversary cannot delete all of them, because of the heavy-edge budget eq. (2.14) for edges out of level 1, and the edge-count budget eq. (2.16) for edges out of level 2 and beyond. At the same time, we try to minimise the number of vertices introduced into the construction so that it will remain $o(n)$ for as large a k as possible.

2.3.4 Level 1, cheapest heavy edges

From each neighbour v of s along the edges just added, add to R the

$$r_1 := \lceil 10,000/\varepsilon^2 \rceil = \Theta(1) \tag{2.20}$$

cheapest *heavy* edges from v to any of the $n' = n(1 - o(1))$ vertices not yet added (see eq. (2.10)), as before also excluding vertex t. Consider this step a *failure* if any added edge has cost greater than $\frac{1}{9}\varepsilon w_0$. For each neighbour v there are n' edges under consideration, with weights i.i.d. $U(0,1)$, and failure occurs iff the number X of edges with weights in the interval $[\frac{1}{11}\varepsilon w_0, \frac{1}{9}\varepsilon w_0]$ is smaller than r_1. Note that $X \sim \mathrm{Bi}(n', (\frac{1}{9} - \frac{1}{11})\varepsilon w_0)$, thus $\mathbb{E}X = (1 - o(1))\,(\frac{1}{9} - \frac{1}{11})\varepsilon s = \Theta(\varepsilon s)$. Failure means that $X < r_1 < \mathbb{E}X/2$, so by Lemma 2.3.2 the probability of failure for a given v is $\leq \exp(-\Theta(\varepsilon s))$. The number of level-1 vertices v is $k + r_0 = \mathcal{O}(s)$, so by the union bound the probability of any failure is

$$\leq \mathcal{O}(s) \exp(-\Theta(\varepsilon s)) \leq \exp(-\Theta(s)), \tag{2.21}$$

by suitable adjustment of the constants implicit in Θ.

This edge selection conditions the costs of the other edges incident on each v, but none will play any role in the analysis. The adversary must leave r_0 edges out of the root, expanding to

$$r_0 r_1 \geq \frac{1}{10}\varepsilon s \cdot 10,000/\varepsilon^2 = 1,000 s/\varepsilon$$

(heavy) edges out of level 1, of which (by eq. (2.14)) he can delete at most $12s/\varepsilon$, leaving (very generously calculated) at least

$$\rho_1 := 120 s/\varepsilon = \Theta(r_0 r_1) \tag{2.22}$$

edges out of level 1. The vertices at the opposite endpoints of these edges constitute level 2.

2.3.5 Level 2, cheapest edges

From each level 2 vertex v in turn, add to R the cheapest r_2 edges to any of the $n' = n(1 - o(1))$ vertices not yet added, again also excluding vertex t from consideration. Here choose

r_2 so as to make

$$\rho_2 := \rho_1 r_2 = 12s^2, \tag{2.23}$$

namely taking

$$r_2 = \frac{12s^2}{\rho_1} = \frac{12s^2}{120s/\varepsilon} = \tfrac{1}{10}\varepsilon s = \Theta(\varepsilon s). \tag{2.24}$$

Consider this step a *failure* if any added edge has cost greater than $\frac{1}{9}\varepsilon w_0$. For each neighbour v there are $n' = (1-o(1))n$ edges under consideration, with weights i.i.d. $U(0,1)$, and failure occurs iff the number X of edges with weights in the interval $[0, \frac{1}{9}\varepsilon w_0]$ is smaller than r_2. Note that $X \sim \mathrm{Bi}(n', \frac{1}{9}\varepsilon w_0)$, thus $\mathbb{E}X = (1-o(1))\frac{1}{9}\varepsilon s = \Theta(\varepsilon s)$. Failure means that $X < r_2 < 0.99\,\mathbb{E}X$, so by Lemma 2.3.2 the probability of failure for a given v is $\leq \exp(-\Theta(\varepsilon s))$. The number of level-2 vertices v is $(k + r_0)r_1 = \mathcal{O}(s)$ so by the union bound the probability of any failure is

$$\leq \exp(-\Theta(\varepsilon s)) \tag{2.25}$$

This edge selection conditions the costs of the other edges incident on each v, but none will play any role in the analysis. The adversary had to leave at least ρ_1 edges out of level 1, expanding to $\rho_1 r_2 = \rho_2 = 12s^2$ edges out of level 2, of which by eq. (2.16) he can delete at most $10s^2$, leaving at least $2s^2$ edges out of level 2. The vertices at the opposite endpoints of these edges constitute level 3.

2.3.6 Level 3, shortest-path trees

We now grow each level-3 vertex v to a tree T_v with d vertices, including v, choosing

$$d := \left\lceil \sqrt{\frac{n \ln n}{2s^3}} \right\rceil < \sqrt{n}. \tag{2.26}$$

We grow these trees one after another, always working within the $n' = n(1-o(1))$ vertices not yet added, and again always excluding vertex t from consideration.

Controlling the lengths of the paths in T_v would allow a choice of d as large as $\sqrt{n}$, but we make it smaller to keep the number of vertices in R as small as possible (and thus keep it to $o(n)$ for s as large as possible).

Here it will be convenient to work with exponentially rather than uniformly distributed edge weights. There are various easy ways to arrange this. We do so by temporarily replacing each uniform weight w with a weight $w' = -\ln(1-w)$; it is standard that these transformed weights are exponentially distributed, and that $w' \geq w$. We construct a shortest-path tree (SPT) of order d using the transformed weights; it will not be an SPT for the original weights, but its paths will cost less under the original weights, which is all that we care about.

Define the distance $\mathrm{dist}(u, v)$ between two vertices to be the cost of a minimum-weight path between them, and define the radius $\mathrm{rad}(T_v)$ of an SPT T_v to be the maximum distance from v to any vertex in T_v. The radius is described by the following claim, which we phrase in a generic setting with n vertices and a root vertex s.

Claim 2.3.4. *In a complete graph K_n with i.i.d. exponential edge weights with mean 1, the radius $X = \mathrm{rad}(T_s)$ of a shortest-path tree T_s of order d is*

$$X = \sum_{i=1}^{d-1} X_i, \tag{2.27}$$

where the X_i are independent random variables with $X_i \sim \mathrm{Exp}(i(n-i))$.

Proof. Following [Jan99], think of the process of finding shortest paths from s to other vertices as first-passage percolation or "infection spreading" starting from s. Let $L := L(r)$ be the set of vertices within radius (distance) r of s; we think of gradually increasing r, starting with $r = 0$ where $L = \{s\}$. It is well known that each edge $(v, u) \in L(r) \times (V \setminus L(r))$ has exponentially distributed weight W conditioned by $W + \mathrm{dist}(s, v) \geq r$, and that these random weights are independent. This can be seen by imagining that infection has spread to radius r from s, including to the vertex v and additionally a length $r - \mathrm{dist}(s, v)$ further along the edge (v, u), and appealing to the memoryless property of the exponential distribution; it can also be verified by analysing Dijkstra's algorithm in this randomised setting.

It follows that the distance X_1 to the vertex nearest s is distributed as $X_1 \sim \mathrm{Exp}(n-1)$; the additional distance to the next vertex is X_2 with $X_2 \sim \mathrm{Exp}(2(n-2))$ and independent of X_1 (for total distance $X_1 + X_2$); and when there are i vertices in the tree, the additional distance to the next is $X_i \sim \mathrm{Exp}(i(n-i))$, with all the X_i independent, for total distance as claimed. ☐

We will only use trees T_v whose radius is $X \leq (1 + \frac{2}{9}\varepsilon)\,\frac{1}{2}\ln n/n < \frac{1}{2}\ln n/n + \frac{1}{9}\varepsilon w_0$. Call a tree a failure (and do not include it in the structure R) if $X > (1 + \frac{2}{9}\varepsilon)\frac{1}{2}\ln n/n$. Declare the construction of level 3 a *failure* if more than $0.01 s^2$ trees fail.

Since eq. (2.27) is monotone increasing in d, the larger the d, the greater the probability of failure, so in the next paragraphs we will pessimistically take d to be $\sqrt{n}$ (ignoring integrality since $\sqrt{n}$ is large). In this case, applying Claim 2.3.4 to T_v, constructed in a complete graph of order $n' = n(1 - o(1))$, the expectation of X is

$$\mu := \mathbb{E}X = \sum_{i=1}^{d-1} \frac{1}{i(n'-i)} = \frac{1+o(1)}{n} \sum_{i=1}^{d-1} \frac{1}{i} = (1+o(1))\frac{\ln d}{n} = (1+o(1))\frac{1}{2}\frac{\ln n}{n}. \tag{2.28}$$

Thus, failure of T_v implies that

$$\frac{X}{\mu} > 1 + \frac{1}{5}\varepsilon. \tag{2.29}$$

To bound the probability of this event we require one more lemma (also used later in proving Lemma 2.4.2).

Lemma 2.3.5 ([Jan18, Theorem 5.1]). *Let $X = \sum_{i=1}^{n} X_i$ with $X_i \sim \operatorname{Exp}(a_i)$ independent rate-a_i random variables, where $a_i \geq 0$. Write $a_* := \min_i a_i$ and $\mu := \mathbb{E}X = \sum_{i=1}^{n} \frac{1}{a_i}$. Then: for any $\lambda = 1 + \varepsilon > 1$,*

$$\mathbb{P}(X \geq \lambda\mu) \leq \lambda^{-1} e^{-a_*\mu(\lambda - 1 - \ln\lambda)} \leq \exp(-\Omega(\alpha_*\mu)) \tag{2.30}$$

for any $\lambda = 1 - \varepsilon < 1$,

$$\mathbb{P}(X \leq \lambda\mu) \leq e^{-a_*\mu(\lambda - 1 - \ln\lambda)} \leq \exp(-\Omega(\alpha_*\mu)), \tag{2.31}$$

and for any $\varepsilon > 0$,

$$\mathbb{P}(|X - \mu| \geq \varepsilon\mu) \leq 2\exp(-\Omega(\alpha_*\mu)). \tag{2.32}$$

The constants implicit in the $\Omega(\cdot)$ expressions are positive and only depend on ε.

Proof. The inequalities in eq. (2.30) and eq. (2.31) in terms of λ are directly from [Jan18, Theorem 5.1]. The remaining expressions, including eq. (2.32), follow immediately. □

From eq. (2.30) of Lemma 2.3.5, the probability of the event in eq. (2.29) (and thus that of T_v failing) is at most

$$\mathbb{P}\left(X - \mu > \frac{1}{5}\varepsilon\mu\right) \leq \exp(-\Omega(\alpha_*\mu)) = \exp(-\Omega(\ln n)), \tag{2.33}$$

using that $a_* = n' = (1 - o(1))n$, μ is given by eq. (2.28), and $\varepsilon = \Theta(1)$.

The total number of trees built is $N = (k + r_0)r_1r_2$, which, with reference to eq. (2.11), eq. (2.18), eq. (2.20), and eq. (2.24), is $\Theta(s^2)$. By eq. (2.33), each tree independently fails with at most some probability $p = o(1)$. Thus, the number of trees surviving dominates $\mathrm{Bi}(N, 1 - p)$, with expectation $\lambda = N(1 - p) = N(1 - o(1))$. Failure at level 3 means that at least $0.01s^2 = \Theta(N)$ trees fail, equivalently the number surviving is at most some $\lambda(1 - \Theta(1))$, which by Lemma 2.3.2 has probability

$$\exp(-\Omega(s^2)). \tag{2.34}$$

Remark 2.3.6. *When construction of a tree T_v rooted at a level-3 vertex v is finished, the edge between any vertex a of T_v and any vertex b in $V' \setminus V(T_v)$ has weight $w(a, b)$ that — still in the uniform model with edge weights temporarily transformed to be exponentially distributed — is exponentially distributed, conditional upon being $\geq \mathrm{rad}(T_v) - d(v, a)$. Equivalently, the edge $\{a, b\}$ gives a v-to-b path (through a) with cost $\mathrm{rad}(T_v) + X_{a,b}$, where the "excess" $X_{a,b}$ has simple exponential distribution $X_{a,b} \sim \mathrm{Exp}(1)$ (with no conditioning). Furthermore, the $X_{a,b}$ are independent, over all choices of a and b.*

Call R_s the now-complete construction on s. Note that there is no conditioning on edges between the remaining vertices; in particular, the SPT infection process (or equivalently Dijkstra's algorithm) as described in Claim 2.3.4 never looked at edges between uninfected vertices.

2.3.7 Symmetric construction on vertex t

Just as we have constructed R_s, we now make a similar construction R_t for vertex t, with the same branching factors out of levels 0, 1, and 2 and similar SPTs on level-3 vertices. Since the number n' of vertices available after constructing R_s still satisfies $n' = (1-o(1))n$, and because the construction on s did not look at nor condition any edge between these vertices, the construction on t enjoys the same properties as that on s.

2.3.8 Edges between the trees on s and t

It remains only to complete paths between s and t, which we do by adding cheap edges (where present) between the SPTs in R_s and those in R_t.

Let T_u be an SPT rooted at a level-3 vertex u of R_s, and T_v one rooted at a level-3 vertex v of R_t. Let a and b be any vertices in T_u and T_v respectively. By Remark 2.3.6, edge $\{a, b\}$ gives a u-to-b path with cost $\mathrm{rad}(T_u) + X_{a,b}$, the collection of all the excesses $X_{a,b}$

being i.i.d. each with distribution $X_{a,b} \sim \text{Exp}(1)$. Thus, $\{a, b\}$ gives a u-to-v path with cost $\leq \text{rad}(T_u) + X_{a,b} + \text{rad}(T_v)$.

Select, and add to the full construction R, any such "middle edge" $\{a, b\}$ having $X_{a,b} \leq \frac{1}{9}\varepsilon w_0$. This completes the construction of R.

2.3.9 Order of R, failure probability, and path costs

It is worth first confirming that the construction uses, as claimed, $o(n)$ vertices. The number of vertices used is of order $(k + r_0)r_1r_2d$, which by eq. (2.18), eq. (2.20), eq. (2.24), and eq. (2.26) is $\mathcal{O}(s^2d)$. Recalling from eq. (2.26) that $d = \left\lceil \sqrt{n \ln n/2s^3} \right\rceil$, as long as the ceiling function does not affect the order of d, the total number of vertices is $\mathcal{O}(s^2d) = \mathcal{O}\left(\sqrt{ns \ln n}\right)$, which is $o(n)$ for $s = o(n/\ln n)$. However, the ceiling function does affect the order of d when $n \ln n/2s^3 < 1$, i.e., when $s > (\frac{1}{2}n \ln n)^{1/3}$; in this case, $d = 1$, the total number of vertices used is $\mathcal{O}(s^2)$, and this is still $o(n)$ if $s = o(\sqrt{n})$. Taking the two cases together, the construction is valid up to any $s = o(\sqrt{n})$, or equivalently for any $k = o(\sqrt{n})$.

Failures at levels 0, 1, and 2 each occur w.p. $\leq \exp(-\Theta(s))$ (by eq. (2.19), eq. (2.21), and eq. (2.25)), and at level 3 w.p. $\leq \exp(-\Omega(s^2))$ (by eq. (2.34)), so by the union bound the probability of any failure is $\leq \exp(-\Theta(s))$.

We now confirm that, assuming that the construction was successful, any s–t path in R *through successful SPTs* has cost $\leq (1 + \varepsilon)w_0$. (Remember that there may be some unsuccessful SPTs.) By assumption of success, any level-0 edge on s or t has cost $\leq \frac{k}{n} + \frac{1}{9}\varepsilon w_0$, any level-1 edge has cost $\leq \frac{1}{9}\varepsilon w_0$, and any level-2 edge also has cost $\leq \frac{1}{9}\varepsilon w_0$. Each successful level-3 tree T in R_s or R_t has radius $\text{rad}(T) \leq (1 + \frac{2}{9}\varepsilon)\frac{1}{2} \ln n/n \leq \frac{1}{2} \ln n/n + \frac{1}{9}\varepsilon w_0$, and each selected middle edge $\{a, b\}$ connects the roots of two trees at an excess cost (above the sum of the two radii) of $X_{a,b} \leq \frac{1}{9}\varepsilon w_0$. The total of the 9 upper bounds in question is

$$2 \cdot \frac{k}{n} + 2 \cdot \frac{1}{2} \ln n/n + 9 \cdot \frac{1}{9}\varepsilon w_0 = (1 + \varepsilon)w_0. \tag{2.35}$$

2.3.10 Robustness of R

We now show that, after the deletion of the k cheapest paths in G, there remains at least one path in R (that uses successful SPTs). Recall from Section 2.3.2 that deletion of the k cheapest paths in G is conservatively modeled as an adversarial deletion subject to: eq. (2.17), the deletion of exactly k edges incident on each of s and t; eq. (2.14), the number of heavy edges deleted at level 1; and eq. (2.16), the total number of edges deleted elsewhere in R (at levels 2 and 3, and joining R_s and R_t).

Without loss of generality we may assume that the adversary does not delete an edge within an SPT T, nor a middle edge from such a tree to a facing one, since deleting the level-2 edge into the level-3 root of T destroys more paths in R at the same budgetary cost.

By the assumption of success, there are at most $0.01s^2$ failed SPTs on each of s and t, and for simplicity we will deal with them by imagining all trees to be successful but allowing the adversary his choice of this many SPTs to delete; by the argument above we can model this as deletion of edges into the roots of these trees, and simply add $0.02s^2$ to this budget.

Let us now allow the adversary to delete k edges from each of s and t, $12s/\varepsilon$ edges out of level 1 for each (double-counting the heavy-edge budget), and $10.02s^2$ edges out of level 2 for each (again double-counting). Can he destroy all s–t paths? We have not yet made any high-probability structural assertion about the middle edges, so this is a probabilistic question: what is the probability, over the randomness still present in the middle edges, that there is an adversarial deletion destroying all paths?

Of the $k+r_0 = \mathcal{O}(s)$ edges on s, the adversary chooses k to delete; there are at most $2^{\mathcal{O}(s)}$ ways to do so. Any choice leaves $\Theta(r_0 r_1) = \Theta(s)$ edges out of level 1, of which the adversary is able to delete a positive fraction, again in at most $2^{\mathcal{O}(s)}$ ways. Any choice leaves $\Theta(s^2)$ edges out of level 2, of which the adversary is able to delete a positive fraction, in at most $2^{\mathcal{O}(s^2)}$ ways. The adversary makes a similar set of choices on t, but still this comes to just $2^{\mathcal{O}(s^2)}$ possible outcomes in all.

A given deletion choice destroys all paths precisely if it leaves no middle edge of excess $\leq \frac{1}{9}\varepsilon w_0$. (Remember that, w.l.o.g., we have excluded deletions in and between the SPTs at level 3.) By construction, any deletion choice leaves $\Theta(s^2)$ edges out of level 2 and thus, by eq. (2.26), $\Theta(s^2 d) = \Omega(\sqrt{ns \ln n})$ vertices in SPTs in each of R_s and R_t, for $\Omega(ns \ln n)$ potential middle edges. A middle edge is *selected* if its excess cost (in the exponential model) is $w' = -\ln(1-w) \leq \frac{1}{9}\varepsilon w_0$, i.e., if $1 - w \geq \exp(-\frac{1}{9}\varepsilon w_0)$, thus is *rejected* with probability $\exp(-\frac{1}{9}\varepsilon w_0)$. There is no path only if every potential edge is rejected, which happens w.p. $\leq \exp(-\frac{1}{9}\varepsilon w_0 \cdot ns \ln n) = \exp(-\Omega(s^2 \ln n))$. Taking the union bound over all adversarial choices, the probability than any choice leaves no paths is

$$2^{\mathcal{O}(s^2)} \exp(-\Omega(s^2 \ln n)) = \exp(-\Omega(s^2 \ln n)). \tag{2.36}$$

This is dominated by the failure probabilities $\exp(-\Theta(s))$ for other steps.

2.3.11 Success for each k, and for all k

We have shown that, for any $k = o(\sqrt{n})$, subject to an absence of failures, we can generate a robust structure $R^{(k)}$ in which, after adversarial deletions, there remains an s–t path of cost $\leq (1+\varepsilon)w_0(k)$. (Remember that w_0 and s are simple functions of k, per eq. (2.11) and eq. (2.12). Here we retain the argument k we usually suppress.) There are two types of failures possible. The first is that the graph fails Lemma 2.3.3's conclusion that "cheap paths are short"; this occurs w.p. $\mathcal{O}(n^{-1.9})$. The second is that $R^{(k)}$ is not robust; this occurs w.p. $\mathcal{O}(\exp(-\Omega(s(k))))$.

Assume success in generating $R^{(k)}$. We claim that $P_1, \dots, P_{k+1}$ all have cost $\leq (1+\varepsilon)w_0(k)$ (call this "cheap"). Suppose not. Then there is some $i \leq k$ for which $P_1, \dots, P_i$ are cheap but P_{i+1} is not. Our adversary's budget allows it to delete $P_1, \dots, P_i$, and by assumption of success this leaves a cheap path P in $R^{(k)}$. Thus there is a cheap $i+1$st path in G, a contradiction.

It follows that *for each* k, $X_{k+1} \leq (1+\varepsilon)w_0(k)$ with probability

$$1 - \mathcal{O}(n^{-1.9}) - \mathcal{O}(\exp(-\Omega(s(k)))). \tag{2.37}$$

A simple calculation shows that w.h.p. $X_{k+1} \leq (1+\varepsilon)w_0(k)$ *simultaneously for all* k in this range, proving the upper bound in eq. (2.3). By the union bound, the probability of failure to build a robust structure for *any* k is at most

$$\begin{aligned}\sum_{k=0}^{\infty} \exp(-\Omega(s(k))) &\leq \ln n \exp(-\Omega(\ln n)) + \sum_{k=\ln n}^{\infty} \exp(-\Omega(k)) \\ &= \exp(-\Omega(\ln n)) = n^{-\Omega(1)}.\end{aligned} \tag{2.38}$$

Including the probability of failure in applying Lemma 2.3.3, the total failure probability is $\mathcal{O}(n^{-1.9} + n^{-\Omega(1)}) = o(1)$.

2.3.12 Limitation to small k

We have established Theorem 2.1.1 up to any $k = o(\sqrt{n})$, and the construction of $R^{(k)}$ was tailored to such values. For levels using heavy edges, fanout is limited to $\mathcal{O}(s)$. On the other hand, the meet-in-the-middle argument requires that each side grow large, to $\Omega(\sqrt{n/s})$. Thus, for small k, a more-than-constant number of levels is needed. Summing heavy edges over this many levels would exceed the target weight $(1+\varepsilon)w_0$, so light edges are needed. The adversary may delete $\Theta(s^2)$ light edges, so the construction must contain at least this many. The construction explicitly required each light edge to lead to a new vertex, and we do not readily see how to do otherwise as long as we are using shortest-path trees, thus intrinsically limiting s (thus k) to $\mathcal{O}(\sqrt{n})$. For larger k, however, we can obtain sufficient heavy-edge fanout in constant depth, permitting a simpler construction described in Section 2.5.

2.4 Edge order statistics

In this section we establish results on order statistics needed in later sections. The kth order statistic of a sample is its kth smallest value. Let $\{W_{(k)}\}_{k=1}^{n-1}$ be the order statistics of a sample of $n-1$ i.i.d. random variables, variously uniform $U(0,1)$ or exponential Exp(1). We choose $n-1$ rather than n as the parameter both because many expressions are more natural in this parametrisation, and because this way $W_{(k)}$ is the cost of the kth cheapest edge incident to a fixed vertex $v \in K_n$.

The following lemma is used in Section 2.6.3.

Lemma 2.4.1. *Let $l = n^{-0.99}$. Consider the unit interval $[0,1]$ with n points placed uniformly and independently at random. Then w.h.p. every interval of length at least $l' \geq l$ contains at least $0.99l'n$ points.*

Proof. Partition the unit interval into contiguous intervals I_i each of length $L := l/1000$, using $\lfloor 1/L \rfloor$ such intervals (possibly leaving a small interval near 1 not covered). Any interval I of length $l' \geq l$ has at least a $998/1000$ fraction of its length covered by intervals $I_i \subset I$, and we will show that w.h.p. every interval I_i contains at least $0.999Ln$ points (that is, at least a 0.999 fraction of the expectation). If so, it follows that I has at least $0.999 \cdot 0.998 l'n \geq 0.99l'n$ points.

The distribution of the number of points in each interval I_i of length L follows the binomial distribution Bi(n, L). By Lemma 2.3.2,

$$\mathbb{P}\left(\text{Bi}(n,L) \leq 0.999Ln\right) \leq \exp(-\Omega(Ln)),$$

where the sign in the Ω is taken as positive. The probability that any interval I_i contains less than 0.999 points is, by the union bound, at most,

$$\lfloor 1/L \rfloor \cdot \exp(-\Omega(Ln)) = \exp(-\Omega(n^{0.01})) = o(1) \tag{2.39}$$

as desired. □

The following lemma is used in eq. (2.73) and eq. (2.88). It shows that the order statistics are simultaneously concentrated around their means.

Lemma 2.4.2. *Let* $\{W_{(k)}\}_{k=1}^{n-1}$ *be the order statistics of* $n-1$ *i.i.d. random variables, either all uniform* $U(0,1)$ *or all exponential* $\mathrm{Exp}(1)$*. For any* $\varepsilon > 0$ *and* $a = a(n) = \omega(1)$*, w.h.p.*

$$1-\varepsilon \le \frac{W_{(k)}}{\mathbb{E}\, W_{(k)}} \le 1+\varepsilon$$

simultaneously for all k *in the range* $a \le k \le n-1$*.*

Proof. Without loss of generality, we may assume that $a \le n/10$.

Exponential case. It is standard that, where $Z_i \sim \mathrm{Exp}(i)$ are independent exponential r.v.s, we may generate the $W_{(k)}$ as

$$W_{(k)} = \sum_{i=1}^{k} Z_{n-i}. \tag{2.40}$$

Using a superscripted E to highlight the exponential model, $W_{(k)}$ has mean

$$\mu_k = \mu_k^{(E)} \coloneqq \mathbb{E}\, W_{(k)} = \sum_{i=1}^{k} \frac{1}{n-i} = H(n-1) - H(n-k-1) \sim \ln(n) - \ln(n-k); \tag{2.41}$$

the change by 1 in the logarithms' arguments avoids $\ln(0)$ when $k = n-1$ and remains asymptotically correct.

By eq. (2.32),

$$\mathbb{P}(|W_{(k)} - \mu_k| \ge \varepsilon\mu) \le 2\exp\left(-\Omega((n-k)\mu_k)\right). \tag{2.42}$$

By the union bound, it suffices to show that the sum over k from a to $n-1$ of the RHS of eq. (2.42) is $o(1)$. We treat the sum in two ranges. For $k \le \frac{n}{2}$, $(n-k)\mu_k \ge \frac{n}{2} \cdot \frac{k}{n} = \frac{k}{2}$. Thus,

$$\sum_{k=a}^{n/2} \exp\left(-\Omega((n-k)\mu_k)\right) \;\le\; \sum_{k=a}^{n/2} \exp\left(-\Omega(k)\right) \;\le\; \mathcal{O}\big(ae^{-\Omega(a)}\big) \to 0, \tag{2.43}$$

since $a = \omega(1)$. For $k > \frac{n}{2}$, for brevity let $\bar{k} = n - k$. Then $\mu_k \sim \ln n - \ln(\bar{k})$ by eq. (2.41) and

$$\sum_{\bar{k}=1}^{n/2} \exp\left(-\Omega((n-k)\mu_k)\right) = \sum_{\bar{k}=1}^{n/2} \exp\left(-\bar{k}\,\Omega(\ln n - \ln(\bar{k}))\right) = \sum_{\bar{k}=1}^{n/2} \left(\frac{\bar{k}}{n}\right)^{\Omega(\bar{k})}$$

$$\leq \left(\frac{1}{n}\right)^{\Omega(1)} \sum_{\bar{k}=1}^{n/2} \bar{k}^{\Omega(1)} \left(\frac{\bar{k}}{n}\right)^{\Omega(\bar{k}-1)} = n^{-\Omega(1)} = o(1), \tag{2.44}$$

where the explicit inequality factors out the $\bar{k} = 1$ term, from which, since $\bar{k}/n \leq 1/2$, the later terms decrease geometrically. This concludes the exponential case.

Uniform case: Let $U_i \sim U(0,1)$ be i.i.d. uniform random variables and $W_i \sim \mathrm{Exp}(1)$ i.i.d. exponential random variables. Because the exponential distribution has CDF $F(x) = 1 - \exp(-x)$, we may couple the two sets of variables as $U_i = F(W_i)$ or equivalently $W_i = f(U_i)$ with $f(x) = F^{-1}(x) = -\ln(1-x)$. Because f is increasing, $W_{(k)} = f(U_{(k)})$. Now using superscript U to distinguish the uniform model, the mean is well known to be

$$\mu_k = \mu_k^{(U)} := \mathbb{E}\, U_{(k)} = \frac{k}{n} \tag{2.45}$$

We want to show that with high probability, for all k in the range $a \leq k \leq n-1$,

$$(1-\varepsilon)\mu_k^{(U)} \leq U_{(k)} \leq (1+\varepsilon)\mu_k^{(U)}$$

or equivalently,

$$f\left((1-\varepsilon)\mu_k^{(U)}\right) \leq W_{(k)} \leq f\left((1+\varepsilon)\mu_k^{(U)}\right).$$

From the exponential case already proved, taking error bound $\varepsilon/2$, we know that w.h.p., for all k,

$$(1-\varepsilon/2)\mu_k^{(E)} \leq W_{(k)} \leq (1+\varepsilon/2)\mu_k^{(E)},$$

so it suffices to show that, for all k (deterministically),

$$f\left((1-\varepsilon)\mu_k^{(U)}\right) \leq (1-\varepsilon/2)\mu_k^{(E)} \quad \text{and} \quad f\left((1+\varepsilon)\mu_k^{(U)}\right) \geq (1+\varepsilon/2)\mu_k^{(E)}.$$

This is so. Using eq. (2.45), eq. (2.41), and convexity of f,

$$f\left((1-\varepsilon)\mu_k^{(U)}\right) = f((1-\varepsilon)k/n) \leq (1-\varepsilon)f(k/n) = (1-\varepsilon)\ln\left(\frac{n}{n-k}\right) \leq (1-\varepsilon/2)\mu_k^{(E)};$$

$$f\left((1+\varepsilon)\mu_k^{(U)}\right) = f((1+\varepsilon)k/n) \geq (1+\varepsilon)f(k/n) = (1+\varepsilon)\ln\left(\frac{n}{n-k}\right) \geq (1+\varepsilon/2)\mu_k^{(E)}.$$

□

2.5 Upper bound for large k, sketch

2.5.1 Introduction

To address larger values of k we use a different construction, generating s–t paths of length 4. A straightforward extension of the previous argument to this construction would let us get up to $k = n - f(n)$ for an arbitrarily slowly growing function f, but not to $k = n - 1$ because it requires $k + 1/\varepsilon^2$ edges incident on each of s and t (thus requires that $k + 1/\varepsilon^2 \leq n - 1$).

Getting all the way to $k = n - 1$ requires a couple of additional ideas. Again, we will introduce an adversary with a cost budget that with high probability exceeds the cost of the first k cheapest paths. First, we observe that much of the adversary's cost budget must be spent on edges incident to s and t, leaving less to delete other edges, thus allowing a smaller structure R to be sufficiently robust. In particular, the k cheapest paths from s to t must use edges incident on s of total weight at least $\sum_{i=1}^{k} W_{(i)}^{s}$ where

$$W_{(i)}^{v} \tag{2.46}$$

is the cost of the ith cheapest edge incident to v. (We may omit the superscript when it is either generic or clear from context.) One technical detail is that, where R includes the $k + r_0$ cheapest edges incident to s, we will control $W_{(k+r_0)} - W_{(k)}$ directly, using results on order statistics from Section 2.4, rather than through a high-probability upper bound on $W_{(k+r_0)}$ and a high-probability lower bound on $W_{(k)}$. Finally, it is no longer adequate to allow path costs to exceed their nominal values by an $\varepsilon = \Theta(1)$ factor, as such large excesses would swell the adversary's budget too quickly, so we more tightly control the excess cost of each path P_k as a function of k (and n, implicitly).

The details later will be clearer if we sketch the argument now, with most details but without the calculations. We will argue for k from $n^{4/10}$ to $n-1$. (We must start with some $k = o(n^{1/2})$ since that is as far as the "small k" argument extended, and we need $k = \omega(n^{1/3})$ since below this the new construction's path costs would exceed the $2k/n$ target.)

2.5.2 Structure R

Figure 2.2 illustrates the robust structure $R = R^{(k)}$ after adversarial deletion of root edges, as discussed in item Section 2.5.6 below. The construction is based on parameters $r_0 = r_0(k)$

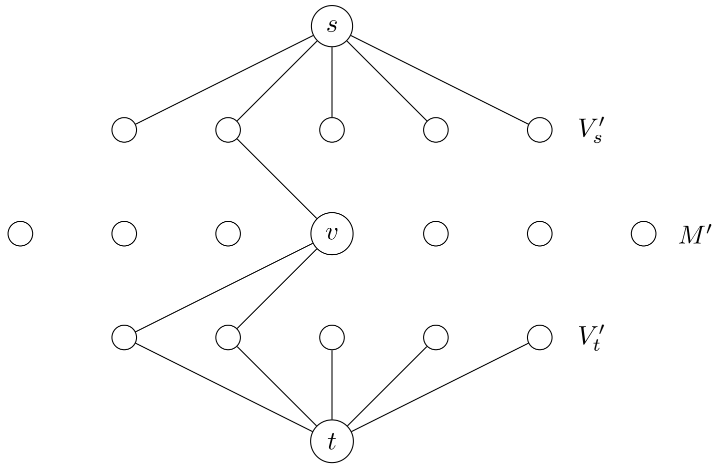


Fig. 2.2 The robust structure $R = R^{(k)}$ after adversarial deletion of k edges on s, leaving r_0 edges to some vertices $V_s' \subseteq V_s$, and likewise for t and V_t'. The middle vertices are pruned to $M' = M \setminus (V_s' \cup V_t')$, and edges from M' to V_s' and V_t' are in R if they have weight between ε_k and $2\varepsilon_k$. Here, edges from just one representative vertex $v \in M'$ are illustrated.

and ε_k to be defined later. Start with R consisting of just the vertices s and t. Add to R the $k + r_0$ edges incident on s of lowest cost, and let V_s be the set of opposite endpoints of these edges. Do the same for t, generating vertex set V_t. Take $M \coloneqq V(G) \setminus \{s,t\}$ as a collection of "middle vertices".

Note that V_s, V_t, and M may well have vertices in common, but our analysis will use a subgraph of R where the relevant subsets of these three sets are disjoint, and it is easier to understand the construction imagining them to be disjoint. Add to R each edge e in $M \times V_s$ and $M \times V_t$ that is "heavy but not too heavy", with cost $W(e) \in (\varepsilon_k, 2\varepsilon_k)$. This concludes the construction of the structure R.

2.5.3 Path weights

It is immediate that every s–t path in R has cost at most

$$W^s_{(k+r_0)} + 2\varepsilon_k + 2\varepsilon_k + W^t_{(k+r_0)}. \tag{2.47}$$

We will show (in eq. (2.74) for uniform and eq. (2.110) for exponential) that, subject to the non-occurrence of certain unlikely failure events, eq. (2.47) is at most

$$W^s_{(k+1)} + W^t_{(k+1)} + 7\varepsilon_k. \tag{2.48}$$

We will show in Section 2.5.6 that, after deletion of the first k paths, there remains an s–t path in R (again subject to non-occurrence of unlikely failure events), whereupon it follows that

$$X_{k+1} \le W^s_{(k+1)} + W^t_{(k+1)} + 7\varepsilon_k. \tag{2.49}$$

2.5.4 Adversary

We define an adversary who is "sufficiently strong" to delete the first k paths. For $k \le n^{4/10}$, taking $\varepsilon = 0.1$, eq. (2.37) implies that w.p.

$$1 - \mathcal{O}\big(n^{-1.9}\big) - \mathcal{O}\big(\exp(-\Omega(n^{4/10}))\big) = 1 - \mathcal{O}\big(n^{-1.9}\big) = 1 - o(1), \tag{2.50}$$

we have that

$$X_k \le X_{n^{4/10}} \le 3n^{4/10}/n. \tag{2.51}$$

For $k > n^{4/10}$, further assume the absence of the failure events alluded to just above, so that eq. (2.49) holds. Then, hiding a sum of the $\ln n/n$ terms of eq. (2.3) in the $o(\,)$ term below,

$$\begin{aligned}
\sum_{i=1}^{k} X_i &= \sum_{i=1}^{n^{4/10}} X_i + \sum_{i=n^{4/10}+1}^{k} X_i \\
&\le \frac{3\big(n^{4/10}\big)^2}{n} + \sum_{i=n^{4/10}+1}^{k} X_i \\
&\le 3n^{-2/10} + \sum_{i=n^{4/10}+1}^{k} \Big(W^s_{(i)} + W^t_{(i)} + 7\varepsilon_{i-1}\Big) \\
&=: U_k.
\end{aligned} \tag{2.52}$$

Thus, the first k paths' edges have total weight at most U_k.

Furthermore, the first k paths' edges incident on s and t are all distinct except, possibly, for the edge $\{s, t\}$. Therefore, not counting edge s–t at all, the cost of these "incident" edges is at least

$$I_k := \sum_{i=1}^{k-1} \left(W^s_{(i)} + W^t_{(i)} \right). \tag{2.53}$$

(In proving Claim 2.8.2 we will use a slightly different lower bound I_k on the weight of the incident edges.)

It follows that the first k paths' "middle edges" (edges other than the incident edges) cost at most $U_k - I_k$. We will explicitly define a budget B_k satisfying

$$B_k \geq U_k - I_k. \tag{2.54}$$

We will allow the adversary to delete any k edges in G incident on each of s and t, possibly including the edge s–t (enough to let it delete the incident edges of the first k paths), and to delete any other edges in G of total cost at most B_k (enough to let it delete the middle edges of the first k paths). Thus, the adversary is sufficiently strong to delete the first k paths.

The adversary's allowable deletions in G mean that also in R it deletes at most k edges incident on each of s and t, and middle edges of total cost at most B_k.

2.5.5 Budgets B_k

The budgets B_k will be defined explicitly in the details. For the model with uniformly distributed edge weights we will do so in two ranges of k, corresponding to Claims 2.6.1 and 2.6.2, and likewise in the model with exponentially distributed edge weights, corresponding to Claims 2.8.1 and 2.8.2. For Claims 2.6.2 and 2.8.2 we will establish eq. (2.54) directly.

For Claims 2.6.1 and 2.8.1 we will establish eq. (2.54) by the following reasoning; we will only need to check eq. (2.55), eq. (2.56), and eq. (2.57) below. We will show that the budgets satisfy

$$B_{k+1} \geq B_k + 8\varepsilon_k. \tag{2.55}$$

(Roughly speaking, given B_k we will set ε_k as small as possible while keeping $R^{(k)}$ robust to the adversary with budget B_k. Then, we will set B_{k+1} as small as possible, namely by taking equality in eq. (2.55). Behind the scenes, we derive B_k by solving the differential-equation equivalent of eq. (2.55) satisfied with equality.)

We will show that eq. (2.54) is satisfied in the base case, by showing that

$$B_k \geq U_k \quad \text{for } k = n^{4/10}. \tag{2.56}$$

Then, eq. (2.54) is established for all k by induction on k:

$$U_{k+1} - I_{k+1} = (U_{k+1} - U_k) - (I_{k+1} - I_k) + (U_k - I_k)$$

which by eq. (2.52), eq. (2.53), and the inductive hypothesis eq. (2.54) is

$$\begin{aligned} &\leq (W^s_{(k+1)} + W^t_{(k+1)} + 7\varepsilon_k) - (W^s_{(k)} + W^t_{(k)}) + B_k \\ &\leq B_k + 8\varepsilon_k \quad \text{(see below)} && (2.57) \\ &\leq B_{k+1} \quad \text{(by eq. (2.55))}. && (2.58) \end{aligned}$$

To justify eq. (2.57) it suffices to show that $W_{(k+1)} - W_{(k)}$ is at most $0.1\varepsilon_k$, and we do so in eq. (2.75) for the uniform case and in eq. (2.112) for the exponential case. In both cases, $r_0 = \omega(1)$, and $W_{(k+r_0)} - W_{(k+1)} = \mathcal{O}(\varepsilon_k)$ (as used in going from eq. (2.47) to eq. (2.48)), making this conclusion unsurprising.[1]

2.5.6 Robustness of R

We wish to make R robust against the adversary, so that after the deletions just described, R should retain an s–t path w.h.p., so that eq. (2.49) holds and X_{k+1} is small. It will suffice to show that, to delete all s–t paths in R,

after deletion of k edges incident on each of s and t, an adversary would still have to delete middle edges of total cost more than B_k, (2.59)

and thus it is powerless to do so.

Obtaining this robustness requires choosing ε_k sufficiently large in the construction. With reference to Fig. 2.2, on deletion of any k edges on each of s and t, the level-1 sets are in effect pruned to V_s' and V_t', each of cardinality r_0. Should V_s' and V_t' have vertices in common, or if $t \in V_s'$ or $s \in V_t'$, then there is an s–t path. So, assume that V_s' and V_t' are disjoint and do not contain s nor t. Consider only middle vertices $M' \subseteq M$ not appearing

[1] In proving Claims 2.6.2 and 2.8.2 we will set $r_0 = 1$, so this reasoning does not apply. Indeed, in Claim 2.8.2 (the large-k exponential case) eq. (2.57) would be false — $W_{(k+1)} - W_{(k)}$ can be much larger than ε_k — but (to reiterate) it is not needed there, as we establish eq. (2.54) directly.

in V_s' nor V_t', i.e., $M' = M \setminus \{V_s' \cup V_t'\}$. We will have $r_0 = o(n)$, so $|M'| = n-2-2r_0 > 0.99n$. Note that edges in $M' \times V_s'$, $M' \times V_t'$, $\{s\} \times V_s'$, and $\{t\} \times V_t'$ are all distinct.

Consider a choice of the k deletions on s and t to be fixed in advance. (We will eventually take a union bound over all such choices.) The weights of edges in $M' \times V_s'$ and $M' \times V_t'$ have not even been observed yet, so each has (unconditioned) $U(0,1)$ distribution, all are independent (by distinctness of the edges), and thus each such edge is included in R with probability ε_k, independently.

A vertex $v \in M'$ is connected to V_s' by

$$Z_v^s \sim \mathrm{Bi}(r_0, \varepsilon_k) \tag{2.60}$$

edges, with mean

$$\lambda := \mathbb{E}\, Z_v^s = r_0 \varepsilon_k. \tag{2.61}$$

Define Z_v^t symmetrically, and note that Z_v^s and Z_v^t are i.i.d. Intuitively, if λ is small, Z_v^s is usually 0, is 1 with probability about λ, and rarely any larger value. So, the probability that v is connected to both V_s' and V_t' is about λ^2, in which case to destroy s–t paths through v the adversary must delete an edge of cost at least ε_k. So, to delete all s–t paths, over the nearly n vertices in M' the adversary would have to delete edges of expected total weight at least

$$\varepsilon_k\, n\, \lambda^2. \tag{2.62}$$

We will choose ε_k so that

$$\varepsilon_k\, n\, \lambda^2 > B_k, \tag{2.63}$$

which hopefully will ensure (see Remark 2.5.1) that a path must remain (i.e., that R is robust).

Let us give a back-of-the-envelope calculation. In the uniform case we expect $W_{(k)}$ to be about k/n, so letting $r_0 = \varepsilon_k n$ means that $W_{(k+r_0)} - W_{(k)}$ will be about ε_k, justifying eq. (2.48). Then eq. (2.61) gives $\lambda = \varepsilon_k^2 n$, so eq. (2.63) indicates that we need to take $\varepsilon_k^5 n^3 > B_k$. As noted after eq. (2.55), roughly speaking, we obtain B_k and ε_k by solving this and eq. (2.55) with equality as a system of differential equations.

Remark 2.5.1. *This intuitive argument proves to be essentially sound, but to make it rigorous will take some work. Chiefly, $\mathbb{P}(Z_v^s > 0)$ is of course not exactly $\mathbb{E}\, Z_v^s = \lambda$ even when λ is small,* *and we will also have to consider the case when λ is large. Also, where the intuition is based on expectations, we must calculate the probability of the "failure" event that all paths can be deleted at a cost less than B_k. Finally, we must take the union bound of this failure event over all choices of root edges at s and t (but, as in the small-k case, this turns out to change nothing).*

2.6 Upper bound for large k, uniform model

In this section we fill in the details of the steps from Section 2.5 and show that they conclude the proof of the upper bound in Theorem 2.1.1. Specifically, to control the *path weights* (these emphasised keywords match section titles) we must show that eq. (2.47) is at most eq. (2.48). For the *adversary* we need only show eq. (2.54); as noted earlier, for large k (Claim 2.6.2) we will do this directly, while for medium k (Claim 2.6.1) we will argue that the *budgets* B_k satisfy eq. (2.56) and eq. (2.57). And for *robustness* we will prove that the probability of failure is small (i.e., it is unlikely that the adversary can destroy all s–t paths in $R^{(k)}$).

2.6.1 Claims, and implications for Theorem 2.1.1

We first state the two precise claims we make for large k, in two ranges. We use symbolic constants C_B, C_ε, C'_B, and C'_ε in the claims and the proofs, as it makes the calculations clearer. Whenever we encounter an inequality that the constants must satisfy, we will highlight with a parenthetical "check" that they do so.

Claim 2.6.1. *For $k \in [n^{4/10}, n - 14\sqrt{n}\,]$, let $B_k = \left(C_B k n^{-3/5} + (2n^{-1/5})^{4/5}\right)^{5/4}$ and $\varepsilon_k = C_\varepsilon n^{-3/5} {B_k}^{1/5}$, with $C_B = 32$ and $C_\varepsilon = 5$. Then, asymptotically almost surely, simultaneously for all k in this range,*

$$X_{k+1} \leq W^s_{(k+1)} + W^t_{(k+1)} + 8\varepsilon_k. \tag{2.64}$$

Remark: In proving Claim 2.6.1 we will set

$$r_0 := \varepsilon_k n. \tag{2.65}$$

From the definitions of B_k and ε_k in Claim 2.6.1, both are increasing in k, and we will make frequent use of the following inequalities. For n sufficiently large,

$$B_k = \Theta\left(k^{5/4}n^{-3/4} + n^{-1/5}\right) \quad (2.66) \qquad \varepsilon_k = \Theta\left(k^{1/4}n^{-3/4} + n^{-16/25}\right) \quad (2.69)$$

$$B_k \le B_n \le 1.01 C_B^{5/4} n^{1/2} \quad (2.67) \qquad \varepsilon_k \le \varepsilon_n \le 1.01 C_\varepsilon C_B^{1/4} n^{-1/2} \quad (2.70)$$

$$B_k \ge B_{n^{4/10}} \ge 2n^{-1/5} \quad (2.68) \qquad \varepsilon_k \ge \varepsilon_{n^{4/10}} \ge 1.14 C_\varepsilon n^{-16/25}. \quad (2.71)$$

Claim 2.6.2. *For $k \in (n - 14\sqrt{n}, n-2]$, let*

$$B_k = C_B'\sqrt{n} \quad \textit{and} \quad \varepsilon_k = C_\varepsilon' n^{-1/6} \tag{2.72}$$

with $C_B' = 78$ and $C_\varepsilon' = 5$. Then, asymptotically almost surely, simultaneously for all k in this range,

$$X_{k+1} \le W^s_{(k+1)} + W^t_{(k+1)} + 8\varepsilon_k.$$

Remark: In proving Claim 2.6.2 we will set $r_0 := 1$. Note that here B_k and ε_k are constants independent of k, but we retain the subscript for consistency with the notation of Section 2.5.1.

We will prove the two claims shortly.

Proof of the upper bound of eq. (2.3) in Theorem 2.1.1. Given $\varepsilon > 0$ from Theorem 2.1.1, apply Lemma 2.4.2 to the order statistics $W^s_{(k)}$ and $W^t_{(k)}$ with ε in the lemma as our $\varepsilon/2$ and $a = n^{4/10}$. Then by Claim 2.6.1 w.h.p., simultaneously for all $k \in [n^{4/10}+1, n - 14\sqrt{n}]$,

$$X_k \le W^s_{(k)} + W^t_{(k)} + 8\varepsilon_{k-1} \le (1+\varepsilon/2)2k/n + 8\varepsilon_{k-1} \le (1+\varepsilon)(2k/n + \ln n/n); \tag{2.73}$$

the key point is that $\varepsilon_{k-1} \le \varepsilon_k = o(k/n)$, which follows from *eq.* (2.69). Specifically, by eq. (2.69), $\varepsilon_k/(k/n) = \mathcal{O}\big(k^{-3/4}n^{1/4} + k^{-1}n^{9/25}\big)$, which by $k \ge n^{4/10}$ is $\mathcal{O}\big(n^{-0.3}n^{0.25} + n^{-4/10}n^{0.36}\big) = o(1)$.

Likewise, by Claim 2.6.2, inequality eq. (2.73) holds w.h.p. simultaneously for all $k \in [n - 14\sqrt{n}, n-2]$. Again, we need only show that $\varepsilon_k = o(k/n)$, which holds because here $k/n = \Theta(1)$ while by definition $\varepsilon_k = o(1)$. □

We now prove the two claims, by filling in the details for Sections 2.5.2 and 2.5.6.

2.6.2 Structure R

With reference to Section 2.5.2, all that we need to confirm is that $k + r_0 \le n-1$. For Claim 2.6.1, by hypothesis $k \le n - 14\sqrt{n}$, and provided that $1.01 C_\varepsilon C_B^{1/4} \le 13$ (check), by eq. (2.70) $\varepsilon_k \le 13n^{-1/2}$, whereupon $r_0 = \varepsilon_k n \le 13\sqrt{n}$. For Claim 2.6.2, with $r_0 = 1$, $k + r_0 \le n - 1$ is immediate.

2.6.3 Path weights

With reference to Section 2.5.3, we establish that the bound eq. (2.48) holds w.h.p. simultaneously for all $k \ge n^{4/10}$. With $W_{(k)}$ representing the cost of the kth cheapest edge incident on some fixed vertex (which we will take to be s and then t in turn), it suffices to show that

$$W_{(k+r_0)} \le W_{(k+1)} + 1.1\varepsilon_k \tag{2.74}$$

holds with high probability for all $k \ge n^{4/10}$.

For Claim 2.6.2, with $r_0 = 1$, eq. (2.74) is immediate. For Claim 2.6.1, with $r_0 = \varepsilon_k n$, generate the variables $W_{(k)}$ by placing $n-1$ points uniformly at random on the unit interval I, associating $W_{(k)}$ with the kth smallest point. It suffices to show that, w.h.p., each interval $(W_{(k+1)}, W_{(k+1)} + 1.1\varepsilon_k)$ contains at least r_0 points. For all $k \in [n^{4/10}, n - 14\sqrt{n} - 1]$, Claim 2.6.1 has $\varepsilon_k \ge n^{-0.99}$ by eq. (2.71), so Lemma 2.4.1 shows that w.p. $1 - \exp(-\Omega(n^{0.01}))$, every interval of length $\ge 1.1\varepsilon_k$ in $[0,1]$ contains at least $r_0 \stackrel{\text{def}}{=} \varepsilon_k n$ points, and in particular this holds for all the intervals $(W_{(k+1)}, W_{(k+1)} + 1.1\varepsilon_k)$.

We assume henceforth that the graph G is "good" in the sense that eq. (2.74) holds for all $k \ge n^{4/10}$ for vertices s and t, and that for all $k \le n^{4/10}$ we have the upper bounds on X_k from eq. (2.3), as proved to hold w.h.p. in Section 2.3.

2.6.4 Adversary

With reference to Section 2.5.4, we need only verify eq. (2.54), and this will be done in the next subsection.

2.6.5 Budgets B_k

With reference to Section 2.5.5, we first establish eq. (2.57). This follows from

$$W^s_{(k+1)} - W^s_{(k)} \le 0.1\varepsilon_k. \tag{2.75}$$

The reasoning for this is the same as for eq. (2.74): each interval of length $0.1\varepsilon_k$ contains at least one point. The parameters are trivial to check.

Next, we show that the parameters of Claim 2.6.1 satisfy eq. (2.54), for which as argued in Section 2.5.5 it suffices to show that they satisfy eq. (2.55) and eq. (2.56). We start with eq. (2.56), the base case. Here $k = n^{4/10}$, $B_k \geq 3n^{-2/10}$ from eq. (2.68), and $U_k = 3n^{-2/10}$ from eq. (2.52), establishing eq. (2.56).

To establish eq. (2.55), first note that $\frac{\partial}{\partial k} B_k = \frac{5}{4} C_B {B_k}^{1/5} n^{-3/5}$ is an increasing function. Then,

$$B_{k+1} - B_k \geq \frac{\partial}{\partial k} B_k = \frac{5}{4} C_B {B_k}^{1/5} n^{-3/5} = \frac{5}{4} \frac{C_B}{C_\varepsilon} \varepsilon_k \geq 8\varepsilon_k,$$

since $C_B \geq \frac{8 \cdot 4}{5} C_\varepsilon$ (check).

We now establish eq. (2.54) for the parameters of Claim 2.6.2. With $k^\star = \lfloor n - 14\sqrt{n} \rfloor$, the point where Claim 2.6.1 ends and just before Claim 2.6.2 begins, the previous case showed that $B_{k^\star} \geq U_{k^\star} - I_{k^\star}$, and by eq. (2.67) $B_{k^\star} \leq 77\sqrt{n}$. Then, for k from $k^\star + 1$ to $n - 2$,

$$\begin{aligned}
U_k - I_k &= (U_{k^\star} - I_{k^\star}) + [(U_k - U_{k^\star}) - (I_k - I_{k^\star})] \\
&\leq B_{k^\star} + \left[\sum_{i=k^\star+1}^{k} (W^s_{(i)} + W^t_{(i)} + 7\varepsilon_{i-1}) - \sum_{i=k^\star}^{k-1} (W^s_{(i)} + W^t_{(i)}) \right] \quad \text{(see eq. (2.52) and eq. (2.53))} \\
&\leq B_{k^\star} + \sum_{i=k^\star+1}^{n-2} 7\varepsilon_{i-1} + (W^s_{(k)} + W^t_{(k)} - W^s_{(k^\star)} - W^t_{(k^\star)}) \\
&\leq 77\sqrt{n} + (14\sqrt{n}) \cdot 7C'_\varepsilon n^{-1/6} + 2 \quad \text{(see eq. (2.72))} \\
&\leq 78\sqrt{n} \\
&\leq B_k \quad \text{(see eq. (2.72))},
\end{aligned} \tag{2.76}$$

using that $C'_B \geq 78$ (check).

2.6.6 Minimum of two binomial variables

Before addressing robustness of the structure R, we require a lemma (Lemma 2.6.4) on the minimum Z of two i.i.d. binomial $\mathrm{Bi}(n, p)$ random variables. There is a genuine difference in the cases when the common mean $\lambda = np$ is large or small: if λ is large then Z is likely to be close to λ, making $\mathbb{E}\, Z = \Theta(\lambda)$; if λ is small then Z will most often be 0, occasionally 1 (with probability about λ^2), and rarely anything larger, making $\mathbb{E}\, Z = \Theta(\lambda^2)$. The lemma relies on the following property of the median of a binomial random variable. (A weaker form of eq. (2.77) and thus of Lemma 2.6.4 can be obtained from Lemma 2.3.2 in lieu of using the median.)

Theorem 2.6.3 (Hamza [Ham95, Theorem 2]). *A binomial random variable X has median satisfying* $|\operatorname{Med}(X) - \mathbb{E}\, X| \leq \ln 2$.

In this discrete setting $\mathrm{Med}(X)$ is not unique: it can be any value m for which $\mathbb{P}(X \leq m) \geq 1/2$ and $\mathbb{P}(X \geq m) \geq 1/2$. [Ham95] defines it uniquely as the smallest integer m such that $\mathbb{P}(X \leq m) > 1/2$; as desired, this gives $\mathbb{P}(X \geq \mathrm{Med}(X)) = 1 - \mathbb{P}(X \leq \mathrm{Med}(X) - 1) \geq 1 - 1/2 = 1/2$. (For other results on the binomial median see Kaas and Buhrman [KB80], in particular, Corollary 1. Stronger results for the Poisson distribution are given by Choi [Cho94], proving a conjecture of Chen and Rubin, and by Adell and Jodrá [AJ05].)

Lemma 2.6.4. *Let Z_1, Z_2 be i.i.d.* $\mathrm{Bi}(n,p)$ *random variables, $Z := \min(Z_1, Z_2)$ and $\lambda := \mathbb{E}\, Z_1 = np$.*

1. *If $\lambda \geq 2$, then*

$$\mathbb{P}(Z \geq 0.65\lambda) > 1/4. \tag{2.77}$$

2. *If $\lambda \leq 2$, then*

$$\mathbb{P}(Z \geq 1) > 0.18\lambda^2. \tag{2.78}$$

Proof. In the first case,

$$\mathrm{Med}(Z_1) \geq \lambda - \ln 2 = \frac{\lambda - \ln 2}{\lambda}\lambda \geq \frac{2 - \ln 2}{2}\lambda \geq 0.65\lambda,$$

so $\mathbb{P}\left(Z_1 \geq 0.65\lambda\right) \geq \mathbb{P}\left(Z_1 \geq \mathrm{Med}(Z_1)\right) \geq 1/2$. The same holds of course for Z_2, and the result follows by independence.

In the second case we again use independence, and here

$$\mathbb{P}\left(Z_1 \geq 1\right) = 1 - (1-p)^n \geq 1 - \exp(-\lambda) = \frac{1 - \exp(-\lambda)}{\lambda} \cdot \lambda \geq 0.43\lambda.$$

The last inequality comes from minimising $\frac{1-\exp(-x)}{x}$ over $0 \leq x \leq 2$; the function is decreasing so the minimum is at $x = 2$. □

2.6.7 Robustness in Claim 2.6.1

With reference to Section 2.5.6, let us complete the robustness argument for Claim 2.6.1, showing that eq. (2.59) holds with high probability. Here we have taken $r_0 = \varepsilon_k n$, so that the number of edges from a middle vertex to V_S' (see eq. (2.60)) is $Z_v^s \sim \mathrm{Bi}(\varepsilon_k n, \varepsilon_k)$, with mean $\lambda = r_0 \varepsilon_k = \varepsilon_k^2 n$ (see eq. (2.61)).

Recall that if λ is small we expect (see eq. (2.62)) that to destroy all paths the adversary will have to delete edges of total weight at least $\varepsilon_k\, n\, \lambda^2 = \varepsilon_k^5 n^3$, which will exceed B_k. And, if λ is large, then each Z_v will have expectation close to $\lambda = \varepsilon_k^2 n$, for a total cost $\varepsilon_k n$ times larger, namely $\varepsilon_k^3 n^2$, and again this exceeds B_k. We now replace these rough calculations with detailed probabilistic ones, applying Lemma 2.6.4 to Z_v in the two cases of λ small and large.

For the adversary to delete all s–t paths via v, he must delete at least

$$Z_v := \min(Z_v^s, Z_v^t)$$

edges, and to destroy all paths he must delete at least

$$N := \sum_{v \in M'} Z_v$$

edges. As described in Section 2.5.6, we imagine a fixed deletion of k edges on each of s and t giving neighbour sets V_s' and V_t' and a set M' of middle vertices; we will eventually take a union bound over all such choices.

If $\lambda \geq 2$, then by Lemma 2.6.4, for each $v \in M'$, $\mathbb{P}(Z_v^s \geq 0.65\lambda) \geq 1/4$. Thus, N stochastically dominates $0.65\lambda \cdot \mathrm{Bi}(0.99n, 1/4)$, with expectation $> 0.1608\lambda n$. We shall consider it a *failure* if $N \leq 0.16\lambda n$. Assuming success, since each edge costs at least ε_k to delete, it costs at least $0.16\varepsilon_k \lambda n = 0.16\varepsilon_k^3 n^2$ to delete them all. This exceeds B_k:

$$\begin{aligned}
\frac{0.16 \cdot \varepsilon_k^3 n^2}{B_k} &= 0.16 \cdot C_\varepsilon^3 n^{-9/5} B_k^{-2/5} n^2 \quad \text{(by definition of } \varepsilon_k\text{)} \\
&\geq 0.15 \cdot C_\varepsilon^3 C_B^{-1/2} n^{1/5} n^{-1/5} \quad \text{(by eq. (2.67))} \\
&> 1,
\end{aligned}$$

using that $0.15 \cdot C_\varepsilon^3 C_B^{-1/2} > 1$ (check).

Failure means that $N/(0.65\lambda) \sim \mathrm{Bi}(0.99n, 1/4) \leq (0.16/0.65)n$. Noting that $0.99 \cdot 1/4 > 0.16/0.65$, by Lemma 2.3.2, the probability of failure is $\exp(-\Omega(n))$. By the union bound, the total of the failure probabilities, over all rounds (values of k) and all adversary choices of the k root edges at s and t, is small:

$$\begin{aligned}\sum_k \binom{k+r_0}{r_0}^2 \cdot \exp(-\Omega(n)) \qquad (2.79)\\ &\leq \sum_k (n^{r_0})^2 \exp(-\Omega(n)) \\ &= \sum_k \exp\left(2\varepsilon_k n \ln n - \Omega(n)\right) \quad (\text{by } r_0 = \varepsilon_k n) \\ &\leq n \exp(-\Omega(n)) \quad (\text{using } \varepsilon_k n = \mathcal{O}\left(n^{1/2}\right) \text{ from eq. (2.70)}) \\ &= o(1).\end{aligned}$$

If $\lambda < 2$, then by Lemma 2.6.4 N stochastically dominates $\mathrm{Bi}(0.99n, 0.18\lambda^2)$, with expectation $> 0.175\lambda^2 n$. We shall consider it a *failure* if $N \leq 0.17\lambda^2 n = 0.17\varepsilon_k^4 n^3$. Each edge costs at least ε_k to delete. Assuming success, it thus costs at least $0.17\varepsilon_k^5 n^3$ to delete them all, which exceeds B_k:

$$\begin{aligned}\frac{0.17\varepsilon_k^5 n^3}{B_k} &= 0.17C_\varepsilon^5 \quad (\text{by definition of } \varepsilon_k) \\ &> 1,\end{aligned}$$

using that $0.17C_\varepsilon^5 > 1$ (check).

By Lemma 2.3.2, the probability of failure is

$$\mathbb{P}\left(N \leq 0.17\varepsilon_k^4 n^3\right) = \exp(-\Omega(\varepsilon_k^4 n^3)). \qquad (2.80)$$

Over all rounds (values of k) and adversary choices of edges incident to s and t, the total failure probability is at most

$$\begin{aligned}\sum_k \binom{k+r_0}{r_0}^2 \cdot \mathbb{P}\left(N < 0.17\varepsilon_k^4 n^2\right) \\ &\leq \sum_k \exp\left(2\varepsilon_k n \ln n - \exp(-\Omega(\varepsilon_k^4 n^3))\right) \\ &\leq n \exp(-\Omega(\varepsilon_k^4 n^3)),\end{aligned}$$

because $\varepsilon_k n \ln n$ is dominated by $\varepsilon_k^4 n^3$: the latter is larger by a factor $\varepsilon_k^3 n^2 / \ln n$, which by eq. (2.71) is $\Omega(n^{-48/25} n^2 / \ln n) = \Omega(n^{2/25} / \ln n) = \omega(1)$. Continuing, this is

$$\begin{aligned} &\leq n \exp(-\Omega(n^{11/25})) \quad \text{(invoking eq. (2.71) again)} \\ &= o(1). \end{aligned} \tag{2.81}$$

2.6.8 Robustness in Claim 2.6.2

Again, our aim is to establish robustness of R by showing that eq. (2.59) holds with high probability, and the argument is similar to but simpler than that of Section 2.6.7.

Since $r_0 = 1$, both V_s' and V_t' have size 1. For a vertex $v \in M'$, let Z_v be the number of paths from V_s' to V_t' via v. There is only one such possible path, hence

$$Z_v \sim \text{Bernoulli}\left(\varepsilon_k^2\right).$$

To destroy all s–t paths the adversary must delete at least

$$N := \sum_{v \in M'} Z_v$$

edges. N stochastically dominates $\text{Bi}(0.99n, \varepsilon_k^2)$, which has expectation $0.99\varepsilon_k^2 n$. We declare the event $N \leq 0.98\varepsilon_k^2 n$ a *failure*. Assuming success, destroying all s–t paths would cost at least $\varepsilon_k N \geq 0.98\varepsilon_k^3 n$. This exceeds B_k, since

$$\frac{0.98\varepsilon_k^3 n}{B_k} = \frac{0.98 C_\varepsilon'^3}{C_B'},$$

and $0.98 C_\varepsilon'^3 > C_B'$ (check).

The probability of failure is

$$\mathbb{P}\left(N \leq 0.98\varepsilon_k^2 n\right) = \exp(-\Omega(\varepsilon_k^2 n)) = \exp(-\Omega(n^{2/3})). \tag{2.82}$$

Over all rounds and adversary choices, using that $\binom{k+r_0}{r_0} = \binom{k+1}{1} \le n$, the total failure probability is at most

$$\begin{aligned} \sum_k \binom{k+r_0}{r_0}^2 &\cdot \mathbb{P}(N \le 0.98\varepsilon^2 n) \\ &\le (14\sqrt{n})\, n^2 \exp(-\Omega(n^{2/3})) \quad \text{(by eq. (2.82))} \\ &= o(1). \end{aligned} \tag{2.83}$$

2.7 Lower bound

In this section, we establish the lower bound in eq. (2.3) of Theorem 2.1.1. Section 2.7.1 establishes the lower bound on X_k directly for $k \le \sqrt{\ln n}$. Values $k \ge \sqrt{\ln n}$ are treated in the subsequent parts. In Section 2.7.2, Lemma 2.7.1 establishes a lower bound on the running totals S_k,

$$S_k \coloneqq \sum_{i=1}^{k} X_i. \tag{2.84}$$

In Section 2.7.3, Lemma 2.7.2 obtains a lower bound on X_k using Lemma 2.7.1's lower bound on S_k, the previously established upper bound on X_k from Theorem 2.1.1, and the monotonicity of X_k.

2.7.1 Lower bound for small k

We begin with $k \le \sqrt{\ln n}$. For any fixed $\varepsilon > 0$, we know from [Jan99] that w.h.p.

$$X_1 > (1-\varepsilon/2)\frac{\ln n}{n}. \tag{2.85}$$

Assuming that eq. (2.85) holds, it follows immediately, and deterministically, that for all $k \le \sqrt{\ln n}$,

$$X_k \ge X_1 \ge (1-\varepsilon/2)\frac{\ln n}{n} \ge (1-\varepsilon)\frac{2k+\ln n}{n}. \tag{2.86}$$

The first inequality holds because the sequence X_k is monotone increasing, the next by assumption on X_1, the next by $k = o(\ln n)$.

2.7.2 Lower bound on the running totals

Lemma 2.7.1. *For any $\varepsilon > 0$, w.h.p., simultaneously for every $k \leq n-1$,*

$$S_k \geq (1-\varepsilon)\sum_{i=1}^{k}\left(\frac{2i+\ln n}{n}\right). \tag{2.87}$$

Proof. Write $W^s_{(i)}$ and $W^t_{(i)}$ for the order statistics of edge weights out of s and t, respectively. By Lemma 2.4.2, w.h.p.,

$$W^s_{(k)}, W^t_{(k)} \in \left[(1-\varepsilon/2)\,\frac{k}{n}, (1+\varepsilon/2)\,\frac{k}{n}\right] \quad \text{for all } k \geq \sqrt[3]{\ln n}, \tag{2.88}$$

and we will assume throughout the proof that eq. (2.88) holds.

We prove the assertion in two ranges of k.

For $\ln^{11/10} n \leq k \leq n-1$, the k paths must use at least $k-1$ edges on each of s and t, all distinct (k edges each, ignoring the edge $\{s,t\}$ if it is used). Then, using eq. (2.88), we get that w.h.p., for all k in the range,

$$\begin{aligned} S_k &\geq \sum_{i=1}^{k-1}\left(W^s_{(i)} + W^t_{(i)}\right) \geq \sum_{i=\sqrt[3]{\ln n}}^{k-1} (1-\varepsilon/2)\frac{2i}{n} \\ &= (1-\varepsilon/2)\left(\sum_{i=1}^{k}\frac{2i+\ln n}{n} - \sum_{i=1}^{k}\frac{\ln n}{n} - \sum_{i=1}^{\sqrt[3]{\ln n}-1}\frac{2i}{n} - \frac{2k}{n}\right) &(2.89) \\ &\geq (1-o(1))(1-\varepsilon/2)\sum_{i=1}^{k}\frac{2i+\ln n}{n} \quad \text{(see below)} &(2.90) \\ &\geq (1-\varepsilon)\sum_{i=1}^{k}\left(\frac{2i+\ln n}{n}\right). &(2.91) \end{aligned}$$

To justify eq. (2.90) it suffices to show that the first sum in eq. (2.89) is of strictly larger order than the other terms. The first sum is at least $\sum_{i=k/2}^{k} 2i/n = \Omega(k^2/n)$, which since $k \geq \ln^{11/10} n$ is also $\Omega(k \ln^{11/10} n/n)$ and $\Omega(\ln^{22/10} n/n)$; we will use all three formulations. The second term is of order $\mathcal{O}(k \ln n/n)$, negligible compared with the middle formulation. The third term is $\mathcal{O}\big(\ln^{2/3} n/n\big)$, negligible compared with the last formulation. And the fourth term, of order $\mathcal{O}(k/n)$, is negligible compared with the first formulation.

For $1 \le k \le \ln^{11/10} n$**,** let $\delta = \varepsilon/3$ and let $G' = G - s - t$. Let N_s and N_t be the endpoints of the cheapest $\ln^3 n$ edges out of s and t respectively. Note that these sets are independent of the edge weights of G'.

If any path P_i, $i \le k$, uses a root edge (edge incident on s or t) *not* amongst the $\ln^3 n$ cheapest edges of s or t, then by eq. (2.88) this edge costs at least $(1-\varepsilon)\ln^3 n/n$, thus $S_k \ge (1-\varepsilon)\ln^3 n/n$. Then eq. (2.87) follows because this is larger than the RHS of eq. (2.87), namely $\Theta((k^2 + k\ln n)/n) = \mathcal{O}\big(\ln^{11/5} n/n\big)$ for this range of k. Thus we may assume that for all $i \le k$, each path P_i goes via some $s' \in N_s$, $t' \in N_t$.

For $s' \in N_s$, $t' \in N_t$, define $A(s',t')$ to be the event that t' is one of the $(n-2)^{1-\delta}$ nearest vertices of s', by cost, in G'. Clearly, for each pair s', t', $\mathbb{P}(A(s',t')) = (n-2)^{-\delta}$. Let A be the union of these events, i.e., the event that any such pair has this property. By the union bound,

$$\mathbb{P}(A) \le \big(\ln^3 n\big)^2 (n-2)^{-\delta} = o(1).$$

We assume henceforth that A does not hold: the $\ln^3 n$ cheapest root edges at s and t do not happen to sample any "nearest" pairs in G'.

Let us temporarily switch from the uniform model (in which we are working) to the *exponential* model G' where the edge weights are i.i.d. Exp(1). By assumption that A does not hold, for each $s' \in N_s$, $t' \in N_t$, the distance $d(s',t')$ stochastically dominates $Y \sim \sum_{i=1}^{n^{1-\delta}} \operatorname{Exp}(i(n-2-i))$ by eq. (2.27). We have $\mathbb{E}\, Y = (1+o(1))(1-\delta)\ln n/n$ by eq. (2.28) (just adjusting its last equation where the value of d is substituted in). Applying Lemma 2.3.5's eq. (2.32) with $\mu = \mathbb{E}\, Y$ as above, $a^\star = n-3$, and $\lambda = 1-\delta$, that in the exponential model G',

$$\mathbb{P}\left(d_{G'}(s',t') \le (1-\delta)\frac{(1+o(1))(1-\delta)\ln n}{n}\right) \le \exp\big(-\Theta(n \cdot \ln n/n \cdot \delta^2)\big) = n^{-\Theta(1)}.$$

Since $(1+o(1))(1-\delta)^2 \ge (1-\frac{3}{4}\varepsilon)$, by the union bound this implies, still in the exponential model G',

$$\mathbb{P}\left(\exists s' \in N_s, t' \in N_t : d_{G'}(s',t') \le (1-\tfrac{3}{4}\varepsilon)\ln n/n\right) \le \big(\ln^3 n\big)^2 n^{-\Theta(1)} = o(1). \quad (2.92)$$

Since $(1-\frac{3}{4}\varepsilon)\ln n/n = o(1)$, by standard coupling arguments (see Remark 2.1.6, eq. (2.7)), this also implies that eq. (2.92) holds in the *uniform* model G in which we are working. (The $1+o(1)$ multiplier present in *eq.* (2.7) can be subsumed into ε, as the argument holds for arbitrary small $\varepsilon > 0$.)

Thus w.h.p., for all $s' \in N_s, t' \in N_t$, we have $d_{G'}(s',t') \ge (1-\frac{3}{4}\varepsilon)\ln n$; assume this holds. We already assumed that each path P_i, $i \le k$, goes via some $s' \in N_s, t' \in N_t$, so its non-root edges contribute at least $d_{G'}(s',t') \geq (1-\frac{3}{4}\varepsilon)\ln n/n$ to S_k. Then, for all k in this range,

$$\begin{aligned} S_k &\geq \sum_{i=1}^{k}(1-\tfrac{3}{4}\varepsilon)\frac{\ln n}{n} + \sum_{i=1}^{k-1}\left(W^s_{(i)} + W^t_{(i)}\right) \\ &\geq \sum_{i=1}^{k}(1-\tfrac{3}{4}\varepsilon)\frac{\ln n}{n} + (1-\tfrac{1}{2}\varepsilon)\sum_{i=\sqrt[3]{\ln n}}^{k-1}\frac{2i}{n} \quad \text{(by eq. (2.88))} \qquad (2.93) \\ &\geq (1-\varepsilon)\sum_{i=1}^{k}\left(\frac{2i+\ln n}{n}\right). \end{aligned}$$

To justify the final inequality, rewrite the second sum in eq. (2.93) as $\sum_{i=1}^{k}\frac{2i}{n} - \frac{2k}{n} - \sum_{i=1}^{\sqrt[3]{\ln n}-1}\frac{2i}{n}$ and observe that both its second term, $2k/n$, and its final term, which is of order $\mathcal{O}\left(\sqrt[3]{\ln n}^{\,2}/n\right)$, are negligible compared with the first sum in eq. (2.93), which is of order $\Omega(k\ln n/n)$. □

2.7.3 Lower bound for large k

Lemma 2.7.2. *For any $\varepsilon > 0$, w.h.p., simultaneously for every $k \in [\sqrt{\ln n}, n-1]$,*

$$X_k \geq (1-\varepsilon)\left(\frac{2k+\ln n}{n}\right).$$

Proof. Let $\delta = \varepsilon^2/9$ and define

$$c_k = \frac{2k+\ln n}{n}, \quad L_k = (1-\delta)\sum_{i=1}^{k} c_i, \quad U_k = (1+\delta)\sum_{i=1}^{k} c_i. \qquad (2.94)$$

W.h.p., simultaneously for all k, $S_k \geq L_k$ (by Lemma 2.7.1) and $S_k \leq U_k$ (by the upper bound of Theorem 2.1.1, already proved). Henceforth, assume that both hold, so $L_k \leq S_k \leq U_k$. The rest of the argument is deterministic. For any positive integer $t < k$, using that X_k is monotone increasing, we have

$$\begin{aligned} tX_k &\geq X_k + \cdots + X_{k-t+1} \\ &= S_k - S_{k-t} \\ &\geq L_k - U_{k-t}. \end{aligned} \qquad (2.95)$$

Thus

$$\begin{aligned}
X_k &\geq \frac{1}{t}(L_k - U_{k-t}) = \frac{1}{t}\left((1-\delta)\sum_{i=1}^{k} c_i - (1+\delta)\sum_{i=1}^{k-t} c_i\right) \\
&\geq \frac{1}{t}\left(\sum_{i=k-t+1}^{k} c_i - 2\delta \sum_{i=1}^{k} c_i\right) \geq \frac{1}{t}(tc_{k-t} - 2\delta k c_k) = c_{k-t} - \frac{2\delta k c_k}{t} \\
&= c_k - \frac{2t}{n} - \frac{2\delta k c_k}{t} \\
&\geq c_k - \frac{tc_k}{k} - \frac{2\delta k}{t} c_k \quad \text{(using that } c_k/k > 2/n\text{)} \\
&= c_k\left(1 - \frac{t}{k} - \frac{2\delta k}{t}\right).
\end{aligned}$$

Ignoring integrality for a moment, setting $t = k\sqrt{2\delta}$ would make the last expression $c_k(1 - 2\sqrt{2\delta})$. Since this $t = \Theta(k) = \omega(1)$, rounding it can be seen to change the expression by a factor $1 + o(1)$, so we may safely write

$$X_k \geq c_k(1 - 3\sqrt{\delta}) = (1-\varepsilon)\frac{2k + \ln n}{n}.$$

□

2.8 Exponential model

In this section we prove Theorem 2.1.2, the analogue of Theorem 2.1.1 for exponentially distributed edge weights.

For small k, results for the exponential case follow from those for the uniform. We first argue that the upper bound of Theorem 2.1.1 also holds in the exponential case for any $k = o(n)$. Couple the two models, so that any edge of weight $w = o(1)$ in one model has cost $w' = w(1 + o(1))$ in the other. The uniform-model upper-bound constructions in Section 2.3 (for $k = o(n^{1/2})$) and Sections 2.5 and 2.6 (for larger k) only use edges of weight $o(1)$ (when $k = o(n)$), and therefore the same upper bounds hold for the exponential model; the multiplicative difference of $1 + o(1)$ can be subsumed into the factor $1 + \varepsilon$ already present. (In the construction of Sections 2.5 and 2.6, the "middle edges" are of cost $o(1)$ for *all* k, but the "incident edges" have larger cost for k large. In particular, for large k, eq. (2.74) will no longer hold in the exponential case until we adjust r_0 and ε_k appropriately.)

For the lower bound too, the argument in Section 2.7 carries over for all $k = o(n)$. The lower bounds L_k on the prefix sums S_k derived in Sections 2.7.1 and 2.7.2 carry over to the exponential case because the edge costs are equal to within $1 + o(1)$ factors in the two models. The upper bounds U_k on the prefix sums are simply the sums of the individual upper bounds on X_k, and we have just argued that these change only by a $1 + o(1)$ factor. Section 2.7.3 only uses L_k and U_k to derive lower bounds on X_k, so with these both changed only by $1 + o(1)$ factors, its results carry over verbatim.

Our task, then, is to prove the upper and lower bounds in Theorem 2.1.2 for larger k. For the upper bound, arguing for $k > n^{0.4}$ (there is no advantage to a larger starting value), we use the same approach as for the uniform model in Section 2.5.

For the lower bound, we argue for $k \geq n^{9/10}$. Unfortunately, the method used in Section 2.7 for the uniform distribution does not extend; let us explain why. The lower bound there came from eq. (2.95), $tX_k \geq L_k - U_{k-t}$, valid for any functions L and U with $L_k \leq S_k \leq U_k$. Here, we would take L_k as the sum I_k of incident edges as in eq. (2.53) and U_k as the sum of the X_k upper bounds as in eq. (2.49). Recall that we defined B_k so that $B_k \geq U_k - L_k$, as in eq. (2.54). Then we can rewrite the previous lower bound approach as $X_k \geq \frac{1}{t}(L_k - U_{k-t}) \geq \frac{1}{t}(L_k - L_{k-t}) + \frac{1}{t}(L_{k-t} - U_{k-t}) \geq \frac{1}{t}\sum_{i=k-t}^{k-1} W_{(i)} - \frac{1}{t}B_{k-t}$. For large k, $W_{(k)}$ and therefore X_k are $\Theta(\ln n)$. Since the B_k grow to size $\Theta(n^{1/2})$ (in the exponential case as well as the uniform case), we are thus limited by the second term to $t = \Omega(n^{1/2+o(1)})$. However, from eq. (2.41), such a large value of t would mean that the average given by the first term is significantly different from $W_{(k)}$.

The desired lower bound would be immediate if we could claim that P_k necessarily used the kth cheapest edge on s (of cost $W^s_{(k)}$) or a later one, and likewise for t. We will prove something close to this. We argue in Section 2.8.7 that every pair of vertices (excluding both s and t) is joined by a path of cost at most δ (for some small δ to be specified) that is edge-disjoint from *all* P_i, $i = 1, \dots, n-1$. We will show that this implies that path P_k uses an edge on s that is at most δ cheaper than $W^s_{(k)}$, and likewise for t, yielding a sufficient lower bound.

2.8.1 Claims, and implications for Theorem 2.1.2

In order to establish upper bounds on X_k in the exponential model, we use the same structure $R^{(k)}$ as described in Section 2.5.2. Then eq. (2.49) follows as before, and we can continue to define U_k as in eq. (2.52). For convenience define

$$\bar{k} = n - k. \tag{2.96}$$

As before we will treat k in two ranges, and we start now with the smaller range.

Claim 2.8.1. *For $k \in [n^{4/10}, n-\sqrt{n}\,]$, let*

$$B_k := \left(\frac{2n^{1/25} + C_B(n^{3/5} - \bar{k}^{3/5})}{n^{1/5}}\right)^{5/4} \quad \text{and} \quad \varepsilon_k := C_\varepsilon B_k^{1/5} n^{-1/5} \bar{k}^{-2/5}, \tag{2.97}$$

with $C_B = 44$ and $C_\varepsilon = 4$. Then, asymptotically almost surely,

$$X_{k+1} \le W^s_{(k+1)} + W^t_{(k+1)} + 8\varepsilon_k. \tag{2.98}$$

Remark: In proving Claim 2.8.1 we will set

$$r_0 := \varepsilon_k \bar{k}. \tag{2.99}$$

because it roughly equates $W_{(k+r_0)} - W_{(k)}$ and ε_k; see eq. (2.41). In this regime integrality is not an issue: r_0 is large, per eq. (2.104).

It is clear that both B_k and ε_k in eq. (2.97) are increasing in k, even over the larger range $k \in [0, n]$. We will make use of the following bounds, holding for n sufficiently large. Here, eq. (2.100) uses that at $k = n - \Theta(\sqrt{n})$, $\bar{k}^{3/5}$ dominates $2n^{1/25}$, while eq. (2.101) takes $k = 0$.

$$B_k \le B_{n-\sqrt{n}} \le C_B^{5/4} n^{1/2} \tag{2.100}$$

$$B_k \ge B_{n^{4/10}} \ge 2n^{-1/5} \tag{2.101}$$

$$\varepsilon_k \le C_\varepsilon B_k^{1/5} n^{-1/5} n^{-1/2 \cdot 2/5} \le C_\varepsilon C_B^{1/4} n^{-3/10} \tag{2.102}$$

$$\varepsilon_k \ge C_\varepsilon (B_{n^{4/10}})^{1/5} n^{-1/5}\, \bar{k}^{-2/5} \ge C_\varepsilon n^{-6/25} \bar{k}^{-2/5} \tag{2.103}$$

$$r_0 = \bar{k}\varepsilon_k \overset{\text{eq. (2.103)}}{\ge} C_\varepsilon n^{-6/25} \bar{k}^{3/5} \ge C_\varepsilon n^{3/50}. \tag{2.104}$$

Claim 2.8.2. *For $k \in (n-\sqrt{n}, n-2\,]$, let*

$$B_k := C'_B \sqrt{n} \quad \text{and} \quad \varepsilon_k := C'_\varepsilon n^{-1/6}, \tag{2.105}$$

with $C_B = 115$ and $C_\varepsilon = 5$. Then, asymptotically almost surely, simultaneously for all k in this range,

$$X_{k+1} \le W^s_{(k+1)} + W^t_{(k+1)} + 8\varepsilon_k. \tag{2.106}$$

Remark: In proving Claim 2.8.2 we will set

$$r_0 := 1. \tag{2.107}$$

As in Claim 2.6.2, B_k and ε_k are constants independent of k, but we retain the subscript for consistency with the notation of Section 2.5.1.

Proof of the upper bounds in Theorem 2.1.2. Analogous to the argument in Section 2.6.1, it is sufficient to check that $\varepsilon_k = o(\mathbb{E}\, W_{(k)})$. Since $\mathbb{E}\, W_{(k)} \sim \ln\left(\frac{n}{n-k}\right) \geq \frac{k}{n}$ (see eq. (2.41)), it is enough to show that $\varepsilon_k = o(k/n)$.

For $k \leq n^{0.99} = o(n)$, by first-order approximation,

$$n^{3/5} - \bar{k}^{3/5} \stackrel{\text{def}}{=} n^{3/5} - (n-k)^{3/5} \sim \frac{3}{5} n^{-2/5} k, \tag{2.108}$$

so $B_k = \Theta\left(n^{-1/5} + n^{-3/4} k^{5/4}\right)$. Hence, from Claim 2.8.1, specifically eq. (2.97),

$$\varepsilon_k = \Theta((n^{-1/25} + n^{-3/20} k^{1/4}) n^{-1/5} n^{-2/5}) = \Theta(n^{-16/25} + n^{-3/4} k^{1/4}) = o(k/n) \tag{2.109}$$

as $k \geq n^{4/10}$.

For $k > n^{0.99}$, we have in Claim 2.8.1 that $\varepsilon_k = \mathcal{O}\left(n^{-3/10}\right)$ by eq. (2.100), and so $\varepsilon_k = o(k/n)$, while in Claim 2.8.2, $\varepsilon_k = \Theta(n^{-1/6}) = o(k/n)$. □

2.8.2 Path weights

To show inequality eq. (2.48) it suffices to show that

$$W_{(k+r_0)} - W_{(k+1)} \leq 1.1 \varepsilon_k. \tag{2.110}$$

In Claim 2.8.2, we have defined $r_0 := 1$, so eq. (2.110) is trivial. For Claim 2.8.1, $\Delta := W_{(k+r_0)} - W_{(k+1)}$ has the same distribution as $\sum_{i=k+2}^{k+r_0} X(n-i)$, where $X(a) \sim \operatorname{Exp}(a)$ and these variables are all independent. Thus Δ is stochastically dominated by the sum of $r_0 - 1$ independent random variables $X(\bar{k} - r_0)$. Since $r_0 = \bar{k}\varepsilon_k$, we have that $\mathbb{E}\, \Delta \leq r_0/(\bar{k} - r_0) = \varepsilon_k/(1-\varepsilon_k)$, and from Lemma 2.3.5 it follows that $\mathbb{P}(\Delta > 1.1\varepsilon_k) = \mathcal{O}(\exp(-\Theta(r_0)))$. From eq. (2.104), by the union bound, there is a negligible chance that eq. (2.48) fails in any round.

2.8.3 Budgets in Claim 2.8.1

As before, we need to define a B_k satisfying eq. (2.54) and, as before, ε_k can be guessed from eq. (2.63), then checked to satisfy yield robustness as in Sections 2.6.7 and 2.6.8. The base case, confirming eq. (2.56), is given by $k = n^{4/10}$, where by eq. (2.101)

$$B_k \geq 2n^{-1/5} \stackrel{\text{def}}{=} U_k. \tag{2.111}$$

To verify eq. (2.55), it is straightforward to check that $\frac{\partial^2}{\partial k^2}B_k$ is positive, so $\frac{\partial}{\partial k}B_k$ is increasing, and

$$B_{k+1} - B_k \geq \frac{\partial}{\partial k}B_k = \frac{5}{4}\frac{3}{5}\frac{C_B}{C_\varepsilon} = \frac{3}{4}\frac{C_B}{C_\varepsilon} \geq 8\varepsilon_k,$$

since $\frac{3}{4}\frac{C_B}{C_\varepsilon} \geq 8$ (check). Finally, we establish eq. (2.57). We show that w.h.p. for all k in the range,

$$\Delta := W^s_{(k+1)} - W^s_{(k)} \leq 0.1\varepsilon_k. \tag{2.112}$$

Note that $\Delta \sim \text{Exp}(\bar{k} - 1)$, so

$$\mathbb{P}\left(\Delta > 0.1\varepsilon_k\right) = \exp\left(-0.1\varepsilon_k \cdot (\bar{k} - 1)\right) = \exp(-\Omega(r_0)) = \exp(-n^{\Omega(1)})$$

by eq. (2.104). Then, by the union bound there is a negligible chance that eq. (2.112) fails for any k.

2.8.4 Robustness in Claim 2.8.1

With reference to Section 2.5.6, we complete the robustness argument for Claim 2.8.1, showing that (2.59) holds with high probability. Here we have taken $r_0 = \varepsilon_k \bar{k}$, so the number of edges from a middle vertex to V'_S (see eq. (2.60)) is $Z^s_v \sim \text{Bi}(\varepsilon_k \bar{k}, \varepsilon_k)$, with mean

$$\lambda = r_0 \varepsilon_k = \varepsilon_k^2 \bar{k} \tag{2.113}$$

(see eq. (2.61)). Recall that if λ is small we expect (see eq. (2.62)) that to destroy all paths the adversary will have to delete edges of total weight at least $\varepsilon_k\, n\, \lambda^2 = \varepsilon_k^5 n \bar{k}^2$, which will exceed B_k. And, if λ is large, then each Z_v will have expectation close to $\lambda = \varepsilon_k^2 \bar{k}$, for a total cost $\varepsilon_k n$ times larger, namely $\varepsilon_k^3 n \bar{k}$, and again this exceeds B_k.

We now show the details of these rough calculations, including the probabilistic details, applying Lemma 2.6.4 to Z_v in the two cases of λ small and large.

For the adversary to delete all s–t paths via v, he must delete at least

$$Z_v := \min(Z_v^s, Z_v^t)$$

edges, and to destroy all paths he must delete at least

$$N := \sum_{v \in M'} Z_v$$

edges. As described in Section 2.5.6, we imagine a fixed deletion of k edges on each of s and t, giving neighbour sets V_s' and V_t' and a set M' of middle vertices, eventually taking a union bound over all such choices.

If $\lambda \geq 2$, then by Lemma 2.6.4, for each $v \in M'$, $\mathbb{P}(Z_v^s \geq 0.65\lambda) \geq 1/4$. Thus, N stochastically dominates $0.65\lambda \cdot \text{Bi}(0.99n, 1/4)$, with expectation $> 0.1608\lambda n$. We shall consider it a *failure* if $N \leq 0.16\lambda n$. Assuming success, since each edge costs at least ε_k to delete, it costs at least $0.16\varepsilon_k \lambda n = 0.16\varepsilon_k^3 n\bar{k}$ to delete them all. This exceeds B_k:

$$\begin{aligned}
\frac{0.16\,\varepsilon_k^3 n\bar{k}}{B_k} &= 0.16\, C_\varepsilon^3 n^{-3/5}\bar{k}^{-6/5} B_k^{-2/5} n\bar{k} \quad \text{(by definition of } \varepsilon_k) \\
&= 0.16\, C_\varepsilon^3 B_k^{-2/5} n^{2/5} \bar{k}^{-1/5} \\
&\geq 0.15\, C_\varepsilon^3 C_B^{-1/2} n^{1/5} n^{-1/5} \quad \text{(by eq. (2.100))} \\
&> 1,
\end{aligned}$$

using that $0.15 \cdot C_\varepsilon^3 C_B^{-1/2} > 1$ (check).

Failure means that $N/(0.65\lambda) \sim \text{Bi}(0.99n, 1/4) \leq (0.16\lambda n)/(0.65n) = (0.16/0.65)n$. Noting that $0.99 \cdot 1/4 > 0.16/0.65$, by Lemma 2.3.2, the probability of failure is $\exp(-\Omega(n))$. By the union bound, the total of the failure probabilities, over all rounds and all adversary choices of the k root edges at s and t, is small:

$$\begin{aligned}
&\sum_k \binom{k + r_0}{r_0}^2 \cdot \exp(-\Omega(n)) \qquad (2.114) \\
&\qquad \leq \sum_k (n^{r_0})^2 \exp(-\Omega(n)) \\
&\qquad = \sum_k \exp\left(2\varepsilon_k \bar{k} \ln n - \Omega(n)\right) \quad \text{(by } r_0 = \varepsilon_k \bar{k}) \\
&\qquad \leq n \exp(-\Omega(n)) = o(1),
\end{aligned}$$

the penultimate inequality using $\varepsilon_k \bar{k} = \mathcal{O}\big(n^{7/10}\big)$ by eq. (2.102).

If $\lambda < 2$, then by Lemma 2.6.4 N stochastically dominates $\mathrm{Bi}(0.99n, 0.18\lambda^2)$, with expectation $> 0.175\lambda^2 n$. We shall consider it a *failure* if $N \leq 0.17\lambda^2 n = 0.17\varepsilon_k^4 n\bar{k}^2$. Each edge costs at least ε_k to delete. Assuming success, it thus costs at least $0.17\varepsilon_k^5 n\bar{k}^2$ to delete them all, which exceeds B_k:

$$\begin{aligned}\frac{0.17\varepsilon_k^5 n\bar{k}^2}{B_k} &= 0.17C_\varepsilon^5 \quad \text{(by definition of } \varepsilon_k\text{)} \\ &> 1,\end{aligned}$$

using that $0.17C_\varepsilon^5 > 1$ (check).

By Lemma 2.3.2, the probability of failure is

$$\mathbb{P}\left(N \leq 0.17\varepsilon_k^4 n\bar{k}^2\right) = \exp(-\Omega(\varepsilon_k^4 n\bar{k}^2)). \tag{2.115}$$

Over all rounds and adversary choices of edges incident to s and t, the total failure probability is at most

$$\begin{aligned}&\sum_k \binom{k+r_0}{r_0}^2 \cdot \mathbb{P}\left(N < 0.17\varepsilon_k^4 n\bar{k}^2\right) \\ &\leq \sum_k \exp\left(2\varepsilon_k \bar{k} \ln n - \exp(-\Omega(\varepsilon_k^4 n\bar{k}^2))\right) \\ &\leq n \exp(-\Omega(\varepsilon_k^4 n\bar{k}^2)),\end{aligned}$$

because $\varepsilon_k^4 n\bar{k}^2$ is larger than $\varepsilon_k \bar{k}$ by a factor $\varepsilon_k^3 n\bar{k}$, which by eq. (2.103) is $\Omega(n^{-18/25}\bar{k}^{-6/5} n\bar{k}) = \Omega(n^{7/25}\bar{k}^{-1/5}) = \Omega(n^{2/25})$. Continuing, this is

$$\begin{aligned}&\leq n \exp(-\Omega(n^{1/25}\,\bar{k}^{2/5})) \quad \text{(invoking eq. (2.103) again)} \\ &= o(1).\end{aligned} \tag{2.116}$$

2.8.5 Budgets in Claim 2.8.2

We now establish eq. (2.54) for the parameters of Claim 2.8.2. Section 2.8.3 showed that eq. (2.54) holds for k up to $k^\star := \lfloor n - \sqrt{n} \rfloor$, the point where Claim 2.8.1 ends and just before Claim 2.8.2 begins, so in particular $B_{k^\star} \geq U_{k^\star} - I_{k^\star}$. For the regime of Claim 2.8.2, we redefine I_k from eq. (2.53). Recall that I_k is a lower bound on the edges incident to s and t used by the first k paths. Previously, the sum defining I_k in eq. (2.53) went to $k-1$ to avoid double counting the $\{s,t\}$ edge. In this regime, however, we need the sum to go

k, as the $W_{(i)}$ increase rapidly. The weight of the $\{s,t\}$ edge is distributed as Exp(1), thus w.h.p. it costs at most $n^{0.01}$. For $k > k^\star$, define

$$I_k \coloneqq \sum_{i=1}^{k} \left(W^s_{(k)} + W^t_{(k)} \right) - n^{0.01}, \tag{2.117}$$

so that w.h.p. I_k is a lower bound on the incident edges: the $n^{0.01}$ term resolves the potential double-counting of $\{s,t\}$. We are now ready to check that eq. (2.54) holds. Following the derivation of eq. (2.76), for k from $k^\star + 1$ to $n-2$,

$$\begin{aligned} U_k - I_k &= (U_{k^\star} - I_{k^\star}) + [(U_k - U_{k^\star}) - (I_k - I_{k^\star})] \\ &\le B_{k^\star} + \sum_{k=k^\star+1}^{n-2} 7\varepsilon_k - (W^s_{(k^\star)} + W^t_{(k^\star)} - n^{0.01}) \quad \text{(see eq. (2.52), eq. (2.53), and eq. (2.117))} \\ &\le 114\sqrt{n} + \sqrt{n} \cdot 7C'_\varepsilon n^{-1/6} + n^{0.01} \quad \text{(see eq. (2.100) and eq. (2.105))} \\ &\le 115\sqrt{n} \\ &\le B_k \quad \text{(see eq. (2.105))}, \end{aligned} \tag{2.118}$$

using that $C'_B \ge 115$ (check).

2.8.6 Robustness in Claim 2.8.2

Again, our aim is to establish robustness of R by showing that eq. (2.59) holds with high probability, and the argument is similar to but simpler than that for robustness in Claim 2.8.1.

Since $r_0 = 1$, both V'_s and V'_t have size 1. For a vertex $v \in M'$, let Z_v be the number of paths from V'_s to V'_t via v. There is only one such possible path, hence

$$Z_v \sim \text{Bernoulli}\left(\varepsilon_k^2\right).$$

To destroy all s–t paths the adversary must delete at least

$$N \coloneqq \sum_{v \in M'} Z_v$$

edges. N stochastically dominates $\text{Bi}(0.99n, \varepsilon_k^2)$, with expectation at least $0.99\varepsilon_k^2$. We declare the event $N \le 0.98\varepsilon_k^2 n$ a *failure*. Assuming success, destroying all s–t paths would cost at least $\varepsilon_k N \geq 0.98\varepsilon_k^3 n$. This exceeds B_k, since by eq. (2.105) and $0.98{C'_\varepsilon}^3 > C'_B$ (check),

$$\frac{0.98\varepsilon_k^3 n}{B_k} = \frac{0.98{C'_\varepsilon}^3}{C'_B} > 1.$$

The probability of failure is

$$\mathbb{P}\left(N \leq 0.98\varepsilon_k^2 n\right) = \exp(-\Omega(\varepsilon_k^2 n)) = \exp(-\Omega(n^{2/3})). \tag{2.119}$$

Over all rounds and adversary choices, using that $\binom{k+r_0}{r_0} = \binom{k+1}{1} \leq n$, the total failure probability is at most

$$\begin{aligned}\sum_k \binom{k+r_0}{r_0}^2 &\cdot \mathbb{P}(N \leq 0.98\varepsilon^2 n) \\ &\leq \sqrt{n}\, n^2 \exp(-\Omega(n^{2/3})) \quad \text{(by eq. (2.119))} \qquad (2.120)\\ &= o(1).\end{aligned}$$

2.8.7 Lower bound

As argued in the introduction of this section, for any $k = o(n)$, the lower bound follows from the uniform case. Thus it is sufficient if we show the lower bound for $k \geq n^{9/10}$, which we do now.

Remark 2.8.3. *With high probability, for every pair of vertices u and v in $G' = G - s - t$, there is a u–v path in G' of cost at most $\delta = 20n^{-1/6}$ that is edge-disjoint from $P_1, \dots, P_{n-1}$.*

Proof. The proof of Claim 2.8.2 showed that w.h.p., for all k in the claim's range (up to $k = n-2$), there is a cheap s–t path (of cost given by eq. (2.106)) disjoint from $P_1, \dots, P_k$, *because* for a given pair of neighbours u, v of s and t, there is a u–v path in G' that is edge-disjoint from these k paths and has cost at most $4\varepsilon_k = 20n^{-1/6} \stackrel{\text{def}}{=} \delta$ (see eq. (2.105)). The existence of a $k+1$st s–t path limits k to $n-2$ since after that there are no new neighbours u and v of s and t, but the rest of the argument extends to $k = n-1$.

In particular, extending the definition eq. (2.105) of B_k and ε_k to $k = n-1$, the derivation of eq. (2.118) extends without change and shows that the budget B_{n-1} covers the middle edges of all paths $P_1, \dots, P_{n-1}$, and the robustness argument also extends and shows eq. (2.119) to hold for $k = n-1$. Since the failure probability in eq. (2.119) is exponentially small, and there are fewer than n^2 pairs $\{u, v\}$ in G', w.h.p. there is a cheap path (of cost $\leq \delta$) for every pair. □

For the remainder of this section we assume that the high-probability conclusion of Remark 2.8.3 holds.

Let H^s_k be the weight of the heaviest edge incident to s used by the first k paths, and let L^s_k be the weight of the lightest edge incident to s *not* used by the first k paths. Define H^t_k and L^t_k likewise.

We claim that for all k from 1 to $n-1$, with $\delta = 20n^{-1/6}$ as in Remark 2.8.3,

$$H^s_k - L^s_k \le \delta. \tag{2.121}$$

We argue by contradiction. Given k, let P_i, $i \le k$, be the path using the edge of weight H^s_k. By Remark 2.8.3, we can construct an s–t path Q whose s-incident edge is the one of weight L^s_k, whose t-incident edge is the same as that of P_i, and whose middle edges cost at most δ and are not used in $P_1, \dots, P_{n-1}$. This path Q is cheaper than P_i: its s-incident edge is cheaper by $H^s_k - L^s_k > \delta$, its t-incident edge has the same cost, and its middle edges (costing at most δ) cost at most δ more than those of P_i. Also, Q is edge-disjoint from the first $i-1$ paths: its s-incident edge L^s_k is not used even by the first k paths, the middle edges are disjoint from those of all $n-1$ paths, and its t-incident edge is that used by P_i (so not used by a previous path). Thus, Q should have been chosen in preference to P_i, a contradiction, establishing eq. (2.121).

Trivially, $H^s_k \ge W^s_{(k)}$. Thus, from eq. (2.121),

$$L^s_k \ge H^s_k - \delta \ge W^s_{(k)} - \delta. \tag{2.122}$$

For $k \le n-2$, the edge of P_{k+1} incident to s costs at least L^s_k and the edge incident to t at least L^t_k. If P_{k+1} is not the single-edge path $\{s,t\}$ these two edges are distinct, so that $X_{k+1} \ge L^s_k + L^t_k$. If P_{k+1} is the single-edge path $\{s,t\}$ then P_k is not, and $X_{k+1} \ge X_k \ge L^s_{k-1} + L^t_{k-1}$. Either way, by eq. (2.122),

$$\begin{aligned} X_{k+1} &\ge L^s_{k-1} + L^t_{k-1} \\ &\ge W^s_{(k-1)} + W^t_{(k-1)} - 2\delta. \end{aligned} \tag{2.123}$$

Recall that we are concerned here with $k \ge n^{9/10}$. By Lemma 2.4.2, for all such k, and for any $\gamma > 0$, w.h.p. $W_{(k)} \ge (1-\gamma)\,\mathbb{E}\,W_{(k)}$. Since the exponential random variable is stochastically greater than the uniform, $\mathbb{E}\,W_{(k)} > k/n = \Omega(n^{-1/10})$, while $\delta = 20n^{-1/6} = o(\mathbb{E}\,W_{(k)})$. From eq. (2.41) it is clear that $\mathbb{E}\,W_{(k-1)} \sim \mathbb{E}\,W_{(k+1)}$ (for any $k = \omega(1)$), and we subsume the asymptotic error into the constant γ. Thus, from eq. (2.123), for any $\gamma > 0$, w.h.p., for all $k \geq n^{9/10}$,

$$X_k \geq (1-\gamma)2\,\mathbb{E}\,W_{(k)},$$

completing the proof of the lower bound in Theorem 2.1.2.

2.9 Expectation

In this section we prove Theorem 2.1.5. We treat the uniform and exponential models at the same time. Let $\mathcal{P}_k$ be the event that P_k exists. Clearly $\mathbb{P}(\mathcal{P}_k) \geq \mathbb{P}(\mathcal{P}_{n-1})$. By Theorem 2.1.1 (for the uniformly random model) and Theorem 2.1.2 (for the exponential model), $\mathbb{P}(\mathcal{P}_{n-1}) = 1 - o(1)$. This establishes the first part of the theorem. Then, let $\mu_k = 2\,\mathbb{E}\,W_{(k)} + \ln n/n$ (so for the uniform model, $\mu_k = w_0(k)$). It suffices to show that

$$\mathbb{E}[X_k \mid \mathcal{P}_k] = (1+o(1))\mu_k \tag{2.124}$$

uniformly in k.

First, we show the lower bound implicit in eq. (2.124). Fix $\varepsilon > 0$. Let $\mathcal{L}_k$ be the event that (jointly) P_k exists and $X_k \geq (1-\varepsilon)\mu_k$. By Theorem 2.1.1 (for the uniform model) and Theorem 2.1.2 (for the exponential model), $\mathcal{L}_k$ holds with probability $1 - o(1)$ uniformly in k. Thus,

$$E[X_k \mid \mathcal{P}_k] \geq \mathbb{P}(\mathcal{L}_k)\,\mathbb{E}[X_k \mid \mathcal{P}_k \wedge \mathcal{L}_k] \geq (1-o(1))\,(1-\varepsilon)\mu_k.$$

Since this holds for any ε, we have that

$$\mathbb{E}[X_k \mid \mathcal{P}_k] \geq (1-o(1))\mu_k.$$

We now establish the corresponding upper bound.

2.9.1 Small k

First, we consider the range $k \leq n^{4/10}$. We will need the following lemma in eq. (2.131).

Lemma 2.9.1. *There exists an absolute constant $C > 0$ such that, for all $\varepsilon > C$, in both the exponential and uniform models, for all $k = o(\sqrt{n})$ the probability of the event*

$$X_k > (1+\varepsilon)\mu_k \tag{2.125}$$

is $\mathcal{O}(n^{-1.9})$.

Proof. By the reasoning given in the introduction of Section 2.8, it is sufficient to show the result in the uniform case, where $\mu_k = \frac{2k+\ln n}{n}$. We use the same argument as developed in Section 2.3, where we prove Theorem 2.1.1 up to $k = o(\sqrt{n})$. Our argument in Section 2.3 (see eq. (2.8)) was that for any sufficiently small $\varepsilon > 0$,

$$\text{if } X_i \leq (1+\varepsilon)\left(\frac{2i}{n} + \frac{\ln n}{n}\right) \text{for all } i \leq k, \text{ then w.h.p. the same holds for } i = k+1. \quad (2.126)$$

We proved this by constructing a structure $R = R^{(k)}$ in G, in which after deleting k paths, each of cost $\leq (1+\varepsilon)(2k/n + \ln n/n)$ from G, w.h.p. there remains a path in R satisfying the same cost bound. By eq. (2.37), the probability of failure was $\mathcal{O}(n^{-1.9}) + \exp(-\Theta(s(k)))$. This does not suffice since for k small the second term may exceed $\mathcal{O}(n^{-1.9})$ (recall $s = 2k + \ln n$).

To prove the lemma, we will show that for some sufficiently *large* constant ε, the failure probability in eq. (2.126) is $\mathcal{O}(n^{-1.9})$. As noted in Remark 2.3.1, a few parts of the argument developed in Section 2.3 rely on ε being sufficiently small, and here we will detail the changes needed. Principally, we will make one modification (a simplification) to Section 2.3's construction of R. We will also track the dependence of key Landau-notation expressions on ε.

Recall from eqs. (2.11) and (2.12) that $s = 2k + \ln n$ and $w_0 = s/n$.

Parallelling the structure of Section 2.3, we start by reviewing the adversary's edge-count budget. This was given by eq. (2.16) which, through its dependence on eq. (2.15), held only for sufficiently small ε. For sufficiently large ε, modulo the one-time failure probability $\mathcal{O}(n^{-1.9})$ from Lemma 2.3.3, each of the first k paths has length $\leq (1+\varepsilon)w_0 \cdot 19n < 20s\varepsilon$, and the total length of the first k paths is at most

$$20ks\varepsilon < 10s^2\varepsilon, \quad (2.127)$$

so we now take this to be the adversary's budget.

We build level-0 edges of R exactly as in Section 2.3.3, and using the same parameter r_0. That is, we add the cheapest $k + r_0$ edges incident on s, with $r_0 = \left\lceil \frac{1}{10}\varepsilon s \right\rceil$ as in eq. (2.18); the opposite endpoints of these edges are the level-1 vertices. Recall that we declared this step a failure if the number X of edges with weights in the interval $[0, \frac{k}{n} + \frac{1}{9}\varepsilon w_0]$ is smaller than $k + r_0$. Note that $X \sim \mathrm{Bi}(n', \frac{k}{n} + \frac{1}{9}\varepsilon w_0)$, thus $\mathbb{E}X = (1 - o(1))\,(k + \frac{1}{9}\varepsilon s)$, and failure means that $X < k + r_0$, i.e., that

$$\frac{X}{\mathbb{E}X} = (1+o(1))\frac{k+\frac{1}{10}\varepsilon s}{k+\frac{1}{9}\varepsilon s} \leq \frac{10}{11}$$

for ε sufficiently large. Then, analogously to eq. (2.19), the failure probability by Lemma 2.3.2 is at most

$$\mathbb{P}(X < \tfrac{10}{11}\,\mathbb{E}X) \leq \exp(-\Omega(\mathbb{E}X)) \leq \exp(-\Omega(\varepsilon s)). \tag{2.128}$$

We skip constructing level-1 edges as in Section 2.3.4, instead setting the level-2 vertices identical to level-1 vertices. (There are no edges between these levels; we have "level 2" only to keep the level numbering the same as before.)

We build level-2 edges exactly as before, with the same parameter r_2, linking to each level-2 vertex its cheapest $r_2 = \frac{1}{10}\varepsilon s$ neighbours (which become the level-3 vertices). The calculations in Section 2.3.5 hold for any $\varepsilon > 0$, and from eq. (2.25) the probability of any failure on this level is

$$\leq \exp -\Theta(\varepsilon s). \tag{2.129}$$

The adversary's deletions of edges incident on s must leave r_0 vertices at level 1 (a.k.a. level 2), thus $r_0 r_2 = \varepsilon^2 s^2/100$ edges leading to level 3. By eq. (2.127) the adversary is allowed to delete at most $10s^2\varepsilon$ edges, so for ε sufficiently large, at least $2s^2$ level-3 vertices remain; this is the same as before, and will continue to suffice.

From level 3 we construct shortest-path trees just as in Section 2.3.6, whose calculations hold for any $\varepsilon > 0$. To recapitulate, these trees are built to a size eq. (2.26) independent of ε, the calculations made are valid for all ε, and the result (here as in Section 2.3) is that each tree fails with some probability $o(1)$, but the level as a whole fails only if at least $0.01s^2$ trees fail, which occurs with probability only $\exp(-\Omega(s^2))$ (see eq. (2.34)).

This concludes the modified construction of R. The remainder of the argument is unchanged from Section 2.3. In the absence of failures, the maximum weight of any s–t path in R remains at most $(1+\varepsilon)w_0$ per eq. (2.35) (indeed, a little less as we've skipped the level-1 edges). The number of successful level-3 trees is $\Omega(s^2)$ as before, and the calculations leading to the probability that an adversary can destroy all cheap paths in R are unaffected: this probability remains $\exp(-\Omega(s^2 \ln n))$ as in eq. (2.36), which is dominated by other failure probabilities.

Tallying up, as in Section 2.3.11, we have a one-time failure probability of $\mathcal{O}(n^{-1.9})$ from Lemma 2.3.3. Out of levels 0, 2 and 3 we have failure probabilities given respectively by eq. (2.128), eq. (2.129) and eq. (2.34), namely $\exp(-\Omega(\varepsilon s))$, $\exp(-\Omega(\varepsilon s))$ and $\exp(-\Omega(s^2))$. Since $s > \ln n$, for some ε sufficiently large, the net failure probability is $\mathcal{O}(n^{-1.9})$, as claimed. □

Let C be the constant in Lemma 2.9.1. Separately, fix any sufficiently small $\varepsilon > 0$. Let

$$\begin{aligned} U_1 &= [0, (1+\varepsilon)\mu_k), \\ U_2 &= [(1+\varepsilon)\mu_k, C\mu_k), \\ U_3 &= [C\mu_k, \infty). \end{aligned}$$

Let $\mathcal{A}_i$ be the event that $X_k \in U_i$. By Theorem 2.1.1, $\mathbb{P}(\mathcal{A}_1) = 1 - o(1)$ and $\mathbb{P}(\mathcal{A}_2) = o(1)$, and by Lemma 2.9.1, $\mathbb{P}(\mathcal{A}_3) = \mathcal{O}(n^{-1.9})$.

Since here we are considering $k \leq n^{4/10} \leq n/2$, with reference to the proof of Remark 2.1.4, one possible choice for P_k is some path of length 2 (there must remain at least one such), and thus, deterministically,

$$X_k \leq W_s + W_t, \tag{2.130}$$

where W_v denotes most expensive edge out of v ($W_v = W^v_{(n-1)}$ in the notation of eq. (2.46)).

In the uniform model, eq. (2.130) means that, deterministically, $X_k \leq 2$. Then,

$$\begin{aligned} \mathbb{E}[X_k] &= \mathbb{P}(\mathcal{A}_1)\,\mathbb{E}[X_k \mid \mathcal{A}_1] + \mathbb{P}(\mathcal{A}_2)\,\mathbb{E}[X_k \mid \mathcal{A}_2] + \mathbb{P}(\mathcal{A}_3)\,\mathbb{E}[X_k \mid \mathcal{A}_3] \\ &\leq (1-o(1)) \cdot (1+\varepsilon)\mu_k + o(1) \cdot (1+C)\mu_k + \mathcal{O}(n^{-1.9}) \cdot 2 \\ &\leq (1+\varepsilon+o(1))\mu_k, \end{aligned} \tag{2.131}$$

since $\mu_k > \ln n/n$. As this holds for arbitrarily small $\varepsilon > 0$,

$$\mathbb{E}[X_k] \leq (1+o(1))\mu_k. \tag{2.132}$$

For the exponential model the same argument applies, once we control $\mathbb{E}[X_k \mid \mathcal{A}_3]$. We make use of the following inequality. Let Z be a random variable with CDF F, and $\mathcal{A}$ be an event with $\mathbb{P}(\mathcal{A}) = \alpha$. Then,

$$\mathbb{E}[Z \mid \mathcal{A}] \leq \mathbb{E}[Z \mid Z > F^{-1}(1-\alpha)]. \tag{2.133}$$

In the case that Z is an exponential random variable with rate λ, $F(z) = 1 - \exp(-\lambda z)$, so $F^{-1}(1-\alpha) = -\ln(\alpha)/\lambda$. By the memoryless property of the exponential, the RHS of eq. (2.133) is $\mathbb{E}[Z] + F^{-1}(1-\alpha)$, giving

$$\mathbb{E}[Z \mid \mathcal{A}] \leq \frac{1-\ln(\alpha)}{\lambda}. \tag{2.134}$$

Recall from eq. (2.40) that $W_v = \sum_{i=1}^{n-1} Z_i$ where $Z_i \sim \mathrm{Exp}(i)$. Condition on the event $\mathcal{A}_3$, taking $\alpha := \mathbb{P}(\mathcal{A}_3) = \mathcal{O}\big(n^{-1.9}\big)$. By eq. (2.134),

$$\mathbb{E}[W_k \mid \mathcal{A}_3] = \sum_{i=1}^{n-1} \mathbb{E}[Z_i \mid \mathcal{A}_3] \leq \sum_{i=1}^{n-1} \frac{1-\ln(\alpha)}{i} \sim (1-\ln(\alpha))\ln n = \mathcal{O}\big(\ln^2 n\big). \tag{2.135}$$

By eq. (2.130), eq. (2.135) and linearity of expectation,

$$\mathbb{P}(\mathcal{A}_3)\,\mathbb{E}[X_k \mid \mathcal{A}_3] \leq \alpha\,\mathbb{E}[W_s + W_t \mid \mathcal{A}_3] = 2\alpha\mathcal{O}\big(\ln^2 n\big) = \mathcal{O}\big(n^{-1.9}\ln^2 n\big), \tag{2.136}$$

which is $o(\mu_k)$ since $\mu_k > \ln n/n$. Thus eq. (2.131) holds also for the exponential model (the change to the middle line of the calculation affects nothing), whereupon so does eq. (2.132).

2.9.2 Large k

For $k \geq n^{4/10}$, we gather the failure events in Section 2.5. First, we have $X_{n^{4/10}} \leq 3n^{4/10}/n$ with failure probability $\mathcal{O}\big(n^{-1.9}\big)$, from eq. (2.51) and eq. (2.50). Then, we have to check two types of failures: failure of eq. (2.48) to be an upper bound on eq. (2.47) (because the edge order statistics are not as expected), and violation of eq. (2.49) (because R fails to be robust against the adversary).

Failure of eq. (2.48) as an upper bound is, in the uniform model, checked through violation of eq. (2.74), the paragraph after eq. (2.74) showing failure to occur w.p. at most $\exp(-\Omega(n^{0.01}))$. Likewise, in the exponential model it is checked in and following eq. (2.110), with a failure probability of $\mathcal{O}\big(\exp(-\Omega(n^{3/50}))\big)$.

The failure probability of eq. (2.49) in the uniform model is calculated for three cases: near eq. (2.79) as $n\exp(-\Omega(n))$, near eq. (2.81) as $n\exp(-\Omega(n^{11/25}))$, and near eq. (2.83) as $14n^{5/2}\exp(-\Omega(n^{2/3}))$. The failure probability in the exponential model is also calculated for three cases: near eq. (2.114) as $n\exp(-\Omega(n))$, near eq. (2.116) as $n\exp(-\Omega(n^{1/25}))$, and near eq. (2.120) $n^{5/2}\exp(-\Omega(n^{2/3}))$.

Thus, the failure probabilities for eq. (2.48) and eq. (2.49) are all $\mathcal{O}(\exp(-n^{0.01}))$, so the probability of any failure affecting any $k > n^{4/10}$ is $\mathcal{O}(n^{-1.9})$.

Let

$$\begin{aligned} U_1 &= [0, (1+\varepsilon)\mu_k) \\ U_2 &= [(1+\varepsilon)\mu_k, \infty), \end{aligned}$$

and let $\mathcal{A}_i$ be the event that P_k exists and $X_k \in U_i$. Thus $\mathbb{P}(\mathcal{A}_1) = 1 - o(1)$ and $\mathbb{P}(\mathcal{A}_2) = \mathcal{O}(n^{-1.9})$.

Conditioning on the event $\mathcal{P}_k$ that P_k exists, this path clearly has cost

$$X_k \le Z := \sum_{v \in V(G)} W_v$$

(analogous to eq. (2.130)). In the uniform model, deterministically, $Z \le n$. In the exponential model, the event $\mathcal{A}_2$ here has the same probability as event $\mathcal{A}_3$ in Section 2.9.1, so we may reuse eq. (2.135), obtaining

$$\mathbb{E}[Z \mid \mathcal{A}_2] = \sum_{v \in V(G)} \mathbb{E}[W_v \mid \mathcal{A}_2] = n\, \mathcal{O}\left(\ln^2 n\right) = o(n^{1.1}).$$

Thus, in both the uniform and exponential cases,

$$\begin{aligned} \mathbb{E}[X_k \mid \mathcal{P}_k] &= \mathbb{P}(\mathcal{A}_1)\, \mathbb{E}[X_k \mid \mathcal{A}_1] + \mathbb{P}(\mathcal{A}_2)\, \mathbb{E}[X_k \mid \mathcal{A}_2] \\ &\le (1 - o(1)) \cdot (1+\varepsilon)\mu_k + \mathcal{O}\left(n^{-1.9}\right) \cdot o(n^{1.1}) \\ &= (1 - o(1))(1+\varepsilon)\mu_k, \end{aligned} \tag{2.137}$$

since $\mu_k > 2k/n > n^{-6/10} = \omega(n^{-0.8})$. As this holds for arbitrarily small $\varepsilon > 0$, for all $k \ge n^{4/10}$,

$$\mathbb{E}[X_k \mid \mathcal{P}_k] \le (1 + o(1))\mu_k, \tag{2.138}$$

completing the proof.

Acknowledgements

We thank Alan Frieze and Wes Pegden for an initial discussion of the second-shortest path. We are also grateful to Alan for noticing that minimum-cost k-flow (Remark 2.1.3) was not an open problem but immediately implied by our other results.

Chapter 3

Bisection width of arbitrary d-regular graphs

3.1 Introduction

Given a graph $G = (V, E)$, a *cut* is a partition of the vertex set V into two disjoint parts (V_0, V_1). A *bisection* is a cut where the two parts are as equal as possible, i.e. if $|V_0|$ and $|V_1|$ differ by at most 1. The *cut size* of a cut is the number of edges $e(V_0, V_1)$ from V_0 to V_1. The *bisection width* $\mathrm{bw}(G)$ of G is the minimum cut size among all bisections of G.

The decision problem of bisection width of arbitrary graphs is known to be NP-complete [GJS76, Theorem 1.3]. It remains NP-complete for d-regular graphs for $d \geq 3$ [BCLS87, Section 2.2], and also for $d = 2$ as in this case G is a union of cycles (and thus $\mathrm{bw}(G) \in \{0, 2\}$): associating weights with the cycle sizes, the hardness follows from the NP-completeness of the partition problem [GJ09, Appendix 3.2].

There are several papers on the bisection width of regular graphs. Kostochka and Mel'nikov [KM92, Theorem 1] show that for all d, one has

$$\mathrm{bw}(G) \leq (d-2)\frac{n}{4} + \mathcal{O}\left(d\sqrt{n}\ln n\right).$$

This was then improved for even d by Monien, Preis, and Diekmann [MPD00] showing that

$$\mathrm{bw}(G) \leq (d-2)\frac{n}{4} + 1$$

for n sufficiently large.

Alon [Alo97] shows

$$\mathrm{bw}(G) \le \left(d - \frac{3\sqrt{d}}{8\sqrt{2}}\right)\frac{n}{4} \tag{3.1}$$

for n sufficiently large, which gives the best asymptotic bound for large d.

For smaller d, Monien and Preis [MP06] show that for $d = 3$ one has $\mathrm{bw}(G) \le n/6 + o(n)$, and for $d = 4$ one has $\mathrm{bw}(G) \le 2n/5 + o(n)$, improving on previous bounds. Stronger bounds can be shown for random d-regular graphs, or for arbitrary d-regular graphs with sufficiently large girth [DSW07; HW16]. As mentioned above, for $d = 2$, G is a union of cycles and thus $\mathrm{bw}(G) \in \{0, 2\}$.

As the graph G could comprise of two components of equal size and therefore have bisection width 0, there are no lower bounds in terms of d that apply to all graphs. However, Bollobás [Bol88, Corollary 2] shows that as $d \to \infty$ almost all d-regular graphs satisfy $\mathrm{bw}(G) \ge \left(d - 2\sqrt{\ln(2)}\sqrt{d}\right)\frac{n}{4}$, showing that eq. (3.1) is optimal up to the constant of the second order term $\sqrt{d}$.

Our main result is as follows.

Theorem 3.1.1. *Let G be a d-regular graph on n vertices. Then,*

$$\mathrm{bw}(G) \le \begin{cases} \frac{1}{6}n + \mathcal{O}(\sqrt{n}) & \text{for } d = 3, \\ \left(d - 3 + \frac{3}{d+1}\right)\frac{n}{4} + d\mathcal{O}(\sqrt{n}\ln n) & \text{for } d \ge 4 \text{ even}, \\ \left(d - 3 + \frac{2(d+3)}{d^2+2d+3}\right)\frac{n}{4} + \mathcal{O}(\sqrt{n}\ln n) & \text{for } d \ge 5 \text{ odd}. \end{cases}$$

For $d = 3$, observe that we obtain the same primary term as in [MP06]. For $d = 4$, we have $(d - 3 + 3/(d+1))/4 = 2/5$, so again we obtain the same leading term, but we have an improved second order term in both cases. We improve the best known bounds for small $d \ge 5$ (up to $d = 125$), after which eq. (3.1) (from [Alo97]) is superior. The improvement is more significant for smaller d, e.g. for $d = 5$ we improve the previous bound of $(d-2)n/4 = 0.75n$ (from [KM92]) to $23n/38 \approx 0.605n$.

The proof for the $d = 3$ case is a simplification of the proof given by [MP06]. The even d case is a simplification and generalization of the $d = 4$ result in [MP06]. Our main contribution is the $d \ge 5$, d odd case, and developing a unifying framework.

3.2 Technical lemmas

In this chapter we describe technical lemmas that we later rely on. It is sufficient if the reader familiarises themself with the statements of Lemmas 3.2.3 and 3.2.4.

Definition 3.2.1. A *weighted token tree* T with maximum weight $p \geq 1$ is a tree such that each vertex v is equipped with $+1$ and -1 tokens, and assigned weight $w(v)$ with $1 \leq w(v) \leq p$. We let $t_+(v), t_-(v) \in \mathbb{N}$ denote the number of $+1$ and -1 tokens at v respectively. We require that the sum of the *number* of tokens and the degree at each vertex does not exceed 3, i.e.

$$t_+(v) + t_-(v) + \deg(v) \leq 3. \tag{3.2}$$

For a subset $S \subset T$ we define the token sum $t(S)$ as the signed *sum* of all tokens in S, i.e.

$$t(S) = \sum_{v \in S} \big(t_+(v) - t_-(v)\big).$$

We denote the total weight of vertices in S as $w(S)$. We define the token density as

$$\delta(S) = \frac{t(S)}{w(S)},$$

so that it may be negative, but we will only deal with weighted token trees with positive density. We write $t = t(T)$, $w = (T)$ and $\delta = \delta(T)$ for short.

For a graph $G = (V, E)$ and a subset $P \subset V$ we define the *external edges* of P with respect to G as the edges between P and $V \setminus P$. We let $\mathrm{ext}_G(P)$ be the number of external edges, i.e.

$$\mathrm{ext}_G(P) := e(P, V \setminus P). \tag{3.3}$$

We use the shorthand $\mathrm{ext}(P)$ if the underlying graph is clear from the context.

The following lemma will be used in proving Lemma 3.2.3, another technical lemma.

Lemma 3.2.2. *Let T be a weighted token tree with maximum weight p, such that $\delta = \delta(T) > 0$ and $p \leq 4/\delta$. Then, there is a set $S \subset V(T)$ with $w(S) \leq 20/\delta$ and*

$$t(S) > \mathrm{ext}(S) = e(S, T \setminus S).$$

The constant 4 in the constraint $p \leq 4/\delta$ is in some sense arbitrary. We may replace it by any larger positive constant at the expense of finding a larger (higher weight) set S.

Proof. We present an algorithm to find a suitable set S. The main idea is as follows. Let $e \in E(T)$ be a (carefully selected) edge and consider forming a new graph T' from T by placing a -1 token on the vertices of e and then removing the edge e. Then a suitable set (for the purposes of the lemma) in either component of T' is suitable in T. (In T we may have not counted an extra edge from S to $T \setminus S$, but in that case we counted an extra -1 token to compensate.)

The algorithm constructs a sequence of trees T_i, staring with $T_0 := T$. We henceforth write $w_i = w(T_i)$ and $t_i = t(T_i)$. The algorithm stops when we first have $w_i \leq 20/\delta$. Each T_i will have a positive token sum and density; see eq. (3.6) and the final paragraph for confirmation.

We describe how to obtain T_{i+1} from T_i. First, we find a vertex $v \in T_i$, such that all (at most 3) subtrees at v (components if we deleted v) have weight at most $w_i/2$. Start with any vertex v and iteratively move towards the subtree that has weight $> w_i/2$, if such exists. We can never move back on an edge to a previous vertex (that would mean both sides of the edge have weight $> w_i/2$). So the algorithm terminates, and when it does we have a suitable v.

Let A be the heaviest (by weight) subtree of v, and let e be the edge from v to A. Let $B := T_i \setminus A$. We have,

$$\frac{w_i}{2} \geq w(A) \geq \frac{w_i - p}{3} \geq \frac{w_i}{3} - \frac{w_i}{12} = \frac{w_i}{4},$$

where the first inequality is by the choice of v, the second by the choice of A (heaviest subtree), and the last inequality follows as $w_i \geq 20/\delta \geq 5p$. As $w_i = w(A) + w(B)$, we also have

$$\frac{3w_i}{4} \geq w_i(B) \geq \frac{w_i}{2}.$$

We let T_{i+1} be the component (A or B) that has the higher density (i.e. "t/w" ratio). If equal, we can pick either. In either case, we have

$$\frac{1}{4} w_i \leq w_{i+1} \leq \frac{3}{4} w_i. \tag{3.4}$$

In T_{i+1}, we place a additional -1 token on the vertex that was part of e. (As we have removed the edge e, there is guaranteed to be place for an extra token, i.e. eq. (3.2) is not violated.) This is a crucial, since a suitable set in T_{i+1} is suitable in T_i, and thus inductively in $T = T_0$. We have

$$t_{i+1} \geq w_{i+1} \frac{t_i}{w_i} - 1, \tag{3.5}$$

where the first term is due to picking the denser component, and the -1 term due to introducing a new -1 token. If $w_{i+1} \leq 20/\delta$, then let $S := T_{i+1}$ as the output. Otherwise, we continue. This concludes the description of the algorithm. It is clear by eq. (3.4) that the algorithm eventually terminates.

Let k be such that the algorithm stopped at T_k. In order to show that T_k is suitable in T, it is sufficient to show that $t_k \geq 1$ (as $\text{ext}_{T_k}(T_k) = 0$, so this implies that T_k is suitable in T_k.). First, we inductively prove that for all $i \leq k$,

$$t_i \geq \delta w_i - 4, \tag{3.6}$$

where $\delta = t/w$. This trivially holds for $i = 0$. Assuming the inductive hypothesis eq. (3.6) up to i, we show that it holds for $i+1$. We have, by eq. (3.5), eq. (3.6) and eq. (3.4),

$$t_{i+1} \geq w_{i+1}\frac{t_i}{w_i} - 1 \geq w_{i+1}\frac{\delta w_i - 4}{w_i} - 1 \geq \delta w_{i+1} - \frac{3}{4}4 - 1 = \delta w_{i+1} - 4. \tag{3.7}$$

The algorithm stopped when we first had $w_k \leq 20/\delta$. Thus $w_{k-1} > 20/\delta$ and so by eq. (3.4) this also implies that $w_k \geq 5/\delta$. Then $t_k \geq \delta \cdot 5/\delta - 4 = 1$ by eq. (3.6) and so T_k is a suitable set in T. □

The following lemma will be used in proving Theorem 3.1.1 for the 3-regular case, specifically in Lemma 3.4.2. The statement is slightly stronger than that of [MP06, Lemma 1], and we draw ideas from their proof in the proof we give below. The proof we give is simpler, but still lengthy.

Lemma 3.2.3. *Let G be a 3-regular graph on n vertices, with each edge coloured black or red. Let R be the number of red edges and let $R = \left(\frac{1}{2} + \varepsilon\right)n$ for some $\varepsilon > 0$. Then, there is a set $X \subset V(G)$ of size $\mathcal{O}(1/\varepsilon)$ such that the number of red edges within X is larger than the number of external black edges, i.e. black edges between X and $V \setminus X$.*

Proof. Let $\varepsilon > 0$. Without loss of generality, ε is sufficiently small. For sets $A, B \subset V(G)$ we define $b(A,B)$ and $r(A,B)$ respectively as the number of black and red edges between vertices in A and B. Let $r(A) = r(A,A)$, that is, the number of red edges within A.

For a set $S \subset V(G)$ we define the *value* of S as

$$v(S) := r(S) - b(S, V \setminus S). \tag{3.8}$$

Our aim is to find a set $X \subset V(G)$ of size $\mathcal{O}(1/\varepsilon)$ with $v(X) \geq 1$.

Let F be the family of connected black components of G, i.e. the connected components of the subgraph induced by the black edges. (A vertex with 3 red edges out is a black component by itself.) For $I \in F$, we have $v(I) = r(I) \geq 0$. We define a connected black component $I \in F$ to be *small* if it has size $\leq 2/\varepsilon$, and *large* otherwise. Let $F_L \subset F$ and $F_S \subset F$ be the family of small and large black components respectively. We have,

$$|F_L| \leq n/(2/\varepsilon) = \varepsilon n/2. \tag{3.9}$$

(By counting the number of black edges, it is easy to see that $\varepsilon n \leq |F_L| + |F_S|$. Thus using eq. (3.9), we obtain that $|F_S| \geq |F_L|$.)

Observe that a small component $I \in F_S$ with at least one internal red edge is suitable and so is the union of two small components connected by a red edge. These have sizes $\mathcal{O}(1/\varepsilon)$, and we henceforth assume no such black components exist. Thus every red edge from a small component has its other endpoint in a large component.

Let us sketch the proof strategy before dwelling into details. For each large black component $I \in F_L$, we let T_I be an arbitrary spanning tree of I on black edges. We build a weighted token tree from each T_I as follows. Place a -1 token at the endpoints for each edge $E(I) \setminus E(T_I)$, that is for each black edge in I but not within the spanning tree T_I. Further, place a $+1$ token at each vertex with a red edge towards a small component. Consider a set in this tree with a higher token sum than external edges. Extending it with the small black components connected to it via red edges will give a set with a positive value. We will use Lemma 3.2.2 to find such a set, and show that it is of suitable size.

Let S and L be the unions of small and large black components respectively. We now establish a lower bound on $r(S, L)$, the number of red edges between small and large components.

For $i \leq 3$, write b_i for the number of vertices with black degree i, and write b_i^S and b_i^L for the number of vertices with black degree i in S and L respectively. Note that $b_0^L = 0$, as vertices without black edges are black components of size 1 and thus small. Hence $b_0 = b_0^S$, and $b_i = b_i^S + b_i^L$.

For a connected black component $I \in F$, let $b_3(I)$ and $b_1(I)$ be the number of vertices within I with degrees 3 and 1 respectively. Further, define the *excess* $d(I)$ as

$$d(I) \coloneqq 2\,(|E(I)| - |V(I)|) = b_3(I) - b_1(I), \tag{3.10}$$

We have $d(I) \geq -2$, equality precisely when I is a tree. Let the *total excess* be

$$d \coloneqq \sum_{I \in F_L} d(I) = b_3^L - b_1^L. \tag{3.11}$$

As $d(I) \geq -2$ for each $I \in F_L$, by eq. (3.9), have

$$d \geq -2|F_L| \geq -\varepsilon n. \tag{3.12}$$

Moreover, as red edges from small components go to large components

$$r(S, L) = 3b_0^S + 2b_1^S + b_2^S. \tag{3.13}$$

Next, by assumption

$$2R = n + 2\varepsilon n = 3b_0 + 2b_1 + 1b_2. \tag{3.14}$$

We have $n = b_3 + b_2 + b_1 + b_0$, thus from eqs. (3.11), (3.13) and (3.14),

$$2\varepsilon n = 2R - n = 2b_0 + b_1 - b_3 \leq 2b_0^S + b_1^S - d \leq \frac{2}{3} r(S, L) - d.$$

Hence,

$$r(S, L) \geq 3\varepsilon n + \frac{3}{2} d. \tag{3.15}$$

Recall Definition 3.2.1 of a weighted token tree. For each $I \in F_L$ let T_I be an arbitrary spanning tree of I on black edges. We define a weighted token tree on each T_I with maximum weight

$$p \coloneqq 1 + 4/\varepsilon \leq 5/\varepsilon.$$

For each $v \in T_I$ place a $+1$ token for each red edge it has going to a small black component, so that the total number of $+1$ tokens placed is $t_+ = r(S, L)$. Write $t_+(v)$ for the number of $+1$ tokens at v and assign weight

$$w(v) = 1 + t_+(v) \cdot (2/\varepsilon). \tag{3.16}$$

to v. We have $w(v) \leq p$, as $t_+(v)$ is at most 2, since v can be joined to at most 2 red edges, since it is in a (large) black component.

Next, for each $I \in F_L$ and for each black edge e in I but not in T_I, place a -1 token at both endpoints of e. The number of -1 tokens in component I is exactly $d(I) + 2$. Let t_- be the total number of -1 tokens, so that by eqs. (3.9) and (3.11),

$$t_- = \sum_{I \in F_L} (d(I) + 2) = d + 2|F_L| \leq d + \varepsilon n. \tag{3.17}$$

Consider a subset $P \subset V(I)$ for some $I \in F_L$ with larger token sum $t(P)$ than edges from P to $V(I) \setminus P$ in T_I, i.e.

$$t(P) > \text{ext}_{T_I}(P). \tag{3.18}$$

Note that by construction (placement of -1 tokens),

$$t_-(P) + \text{ext}_{T_I}(P) \geq \text{ext}_I(P). \tag{3.19}$$

Let Q be the union of P together with all small black components connected to P via red edges in G. Then, by construction (eq. (3.16))

$$|Q| \leq w(P)$$

as for each $+1$ token the weight of v was increased by $2/\varepsilon$ and we include a small black component of size $\leq 2/\varepsilon$. Consider the value $v(Q)$. We have $r(Q) = t_+(P)$ and $b(Q, V \setminus Q) = \text{ext}_I(P)$. Thus, from eq. (3.19),

$$v(Q) = t_+(P) - \text{ext}_I(P) \geq t(P) - \text{ext}_{T_I}(P) \geq 1.$$

Thus it is sufficient to find a set P as above of weight $w(P) = \mathcal{O}(1/\varepsilon)$ that satisfies eq. (3.18) to finish the proof of the lemma. We find such a set using Lemma 3.2.2.

Let $t = \sum_{I \in F_L} t(T_I)$, $w = \sum_{I \in F_L} w(T_I)$ and $\delta = t/w$. We have that,

$$t = t_+ - t_- \overset{\text{eq. (3.17)}}{\geq} r(S, L) - d - \varepsilon n \overset{\text{eq. (3.15)}}{\geq} d/2 + 2\varepsilon n \overset{\text{eq. (3.12)}}{\geq} \varepsilon n. \tag{3.20}$$

Using eq. (3.17) and penultimate expression above, we have that $t_- \leq 2t$ and so $t_+ \leq 3t$, which in turn gives that,

$$w \leq n + t_+/\varepsilon \leq n + 3t/\varepsilon. \tag{3.21}$$

We can now give a bound on the token density. The last inequality will follow as the penultimate expression is increasing in t, thus we can substitute in the last expression from eq. (3.20) to obtain a lower bound. Thus,

$$\delta = \frac{t}{w} \geq \frac{t}{n + 3t/\varepsilon} \geq \frac{\varepsilon}{4}.$$

As δ is just an 'average', we can find a tree $T := T_I$ for some $I \in F_L$ that has

$$\delta(T) \geq \delta \geq \frac{\varepsilon}{4}.$$

Further,

$$\delta(T) \leq \frac{t_+(T)}{t_+(T) \cdot 2/\varepsilon} = \varepsilon/2 \leq 4/p.$$

Thus we can apply Lemma 3.2.2 to T to obtain a set of weight $\mathcal{O}(1/\varepsilon)$ that has a higher token sum than external black edges. Thus by the remark above we can find a suitable set in G whose size is at most the weight of the set found, completing the proof. □

The following lemma is used in Sections 3.5 and 3.6, which contain the proof for the cases $d \geq 4$. Specifically, we use it to obtain eq. (3.38) and eq. (3.51). The lemma is a generalization of [MP06, Lemma 2].

Lemma 3.2.4. *Let $\beta > 0$ and $G = (V, E)$ be a graph on n vertices with maximum vertex degree at most $\Delta \geq 3$ and*

$$\frac{|E|}{|V|} = 1 + (\Delta - 2)\beta. \tag{3.22}$$

Then, there exists an induced subgraph $S \subset G$ such that it has at least $|S| + 1$ internal edges and $|S| \leq \mathcal{O}\left(\frac{\ln n}{\beta}\right)$ uniformly in β (i.e. the constant implicit in $\mathcal{O}(.)$ does not depend on β).

Proof. Observe the following.

(W1) Without loss of generality

$$\beta > \frac{\ln n}{n}, \tag{3.23}$$

or else $S := G$ satisfies the lemma.

(W2) Without loss of generality there are no vertices of degree 1, as deleting them increases the edge to vertex ratio ($|E|/|V|$) and thus β, and we apply the lemma again to a smaller graph. A suitable set S in the smaller graph is trivially suitable in the original graph G.

(W3) Without loss of generality there are no paths of length $> 1/\beta$ on degree 2 vertices. Deleting a path of length $k > 1/\beta$ would remove $k+1$ edges and k vertices, a ratio of $\frac{k+1}{k} = 1 + \frac{1}{k} < 1 + \beta$, thus in the new graph we have a higher edge to vertex ratio ($|E|/|V|$) and thus a higher implied β. Again, we apply the lemma to the new smaller graph to get a suitable set in the original graph.

(W4) Without loss of generality G is connected, as we can apply the lemma to the component with the highest $\frac{|E|}{|V|}$ ratio, for which it will be higher (or equal) than it is for G.

Let d_i be the number of vertices of degree i in G. We have $d_1 = 0$ by assumption **(W2)**. From eq. (3.22), we have that,

$$2(\Delta - 2)\beta n = 2(|E| - |V|) = \sum_{i=2}^{\Delta}(i-2)d_i \leq (\Delta - 2)(n - d_2)$$

$$2\beta n \leq n - d_2. \tag{3.24}$$

The number of degree 3 or higher vertices in G by eqs. (3.23) and (3.24) is

$$d_{\geq 3}(G) = n - d_2 \geq 2\beta n \geq 2 \ln n. \tag{3.25}$$

Next, we form a multigraph F by contracting maximal paths on degree 2 vertices (connecting degree 3 or higher vertices) into a single edge. Each contracted path has length $\leq 1/\beta$ by assumption **(W3)**. Every vertex in F has degree at least 3. F is non-empty by eq. (3.25).

Consider a breadth first search (BFS) on F. Stop when we find the first cycle C. Then C has size at most $2\log_2 n$, since every vertex has degree ≥ 3 in F.

Consider a BFS again starting from the whole of C. Stop when we either obtain a new cycle joined by a path to C, or a path leading back to C (including the case when the path is a chord in C). Again, the total length of any of these is at most $2\log_2 n$. Let this new set of vertices be R. Note that R has at least $|R| + 1$ internal edges and that $|R| = \mathcal{O}(\ln n)$.

The edges in R represent either edges in G or contracted paths of length $\leq 1/\beta$. Expand the contracted paths back in into the original graph, and we obtain a set S of size $\mathcal{O}(\ln n/\beta)$. This set S now fulfils the lemma. □

3.3 Method overview

In this chapter we describe the main methods used.

3.3.1 Helpful sets and proof approach

Definition 3.3.1. For a cut (V_0, V_1) of a graph G and a set $S \subset V_0$ we define the *helpfulness* of S to be

$$h(S) = e(V_0, V_1) - e(V_0 \setminus S, V_1 \dot{\cup} S) = e(S, V_1) - e(S, V_0 \setminus S)$$

i. e. $h(S)$ is the decrease in the cut size if we move S from V_0 to V_1. S is said to be *helpful* if $h(S) > 0$.

When the underlying graph is not clear from the context we will denote the helpfulness as $h_G(S)$.

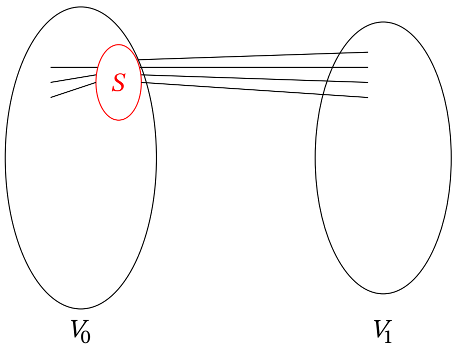


Fig. 3.1 A helpful set S of helpfulness $h(S) = 4 - 3 = 1$.

In other words, a set $S \subset V_0$ is helpful if moving the whole set to V_1 decreases the cut size. For a set $S \subset V_0$, we define the *marginal helpfulness* of a vertex v with respect to S as

$$h(v, S) = \begin{cases} e(v, V_1) - e(v, V_0 \setminus S) & \text{if } v \in S, \\ 0 & \text{if } v \notin S, \end{cases} \tag{3.26}$$

so that $h(S) = \sum_{v \in S} h(v, S) = \sum_{v \in V_0} h(v, S)$.

Remark 3.3.2. *Observe the following:*

$$h(S) = \sum_{v \in S} \big(e(v, V_1) - e(v, V_0)\big) + 2e(S), \tag{3.27}$$

where $e(S)$ is the number of edges within $G[S]$.

Example 3.3.3. A vertex $v \in V_0$ with $e(v, V_1) > e(v, V_0)$ is a *helpful* set by itself.

Example 3.3.4. For d even, $d = 2h$, a set $S \subset V_0$ where each $v \in S$ is incident to exactly $h-1$ cut edges (and so exactly $h+1$ vertices in V_0) is *helpful* if and only if it has $|S|+1$ or more internal edges by eq. (3.27).

Our approach in establishing the claimed bounds on the bisection width in Theorem 3.1.1 is as follows. We start with any bisection (V_0, V_1) of G. At each step, relabelling if necessary so that $|V_0| \geq |V_1|$, we find a helpful set $S \subset V_0$, and move it to V_1, decreasing the cut size. We repeat this step while we can find a helpful set that is *sufficiently small*, ensuring that the cut stays roughly balanced, as helpful set is always moved from the larger part to the smaller one. We stop when we can no longer find a small S, but this will mean the cut size is small.

Finally, we obtain a bisection from the current cut by naively moving vertices from the larger side to the smaller one. Since the cut was kept roughly balanced, the cut size will not increase too much.

3.3.2 Reversible transformations

We introduce the concept of reversible transformations. Reversible transformations transform a graph G to a new graph G' on the same vertex set, such that helpful sets in G' translate to helpful sets in G. This allows us continue to look for a helpful set in G', while gaining desirable properties of the structure of G'.

Definition 3.3.5 (Reversible transformations)**.** Let $G = (V, E)$ be a d-regular graph with a cut (V_0, V_1). A *reversible transformation* from G to $G' = (V, E')$ is obtained by deleting and adding edges to G, such that the resulting graph G' is d-regular.

The transformation comes equipped with: a function $r : \mathcal{P}(V_0) \to \mathcal{P}(V_0)$, that transforms helpful sets from G' to G; and sets $I, R \subset V_0$. Write h for helpfulness in G and h' for helpfulness in G' with respect to the cut (V_0, V_1).

For all $S' \subset V_0$, writing $S = r(S')$ for the translation of S from G' to G, we require the following properties to hold:

(P1) $h(S) \geq h'(S')$,

(P2) $S' \subset S \subset S' \cup I$,

(P3) $S = S'$ if $S' \cap R = \varnothing$,

(P4) edges may only be added or deleted from G if at least one its endpoints lies in I or R.

In other words, I controls which vertices S' may be enlarged with during the translation, and R controls when we need to change the set S'. When translating a helpful set from G' back to G, its size may grow by at most $|I|$ (by property **(P2)**.

The first example of a reversible transformation is given in the proof of Lemma 3.4.1.

Remark 3.3.6. *In a reversible transformation, since G and G' are both d-regular, each vertex has exactly the same number of edges deleted from it as new ones added to it. For sets $S \subset V_0$ with $R \cup I \subset S$, by property **(P4)** this implies that $h(S) = h'(S)$.*

Further, for the reversible transformations that we will define, it will always hold that

$$h'(S') \leq h'(S) \leq h(S), \tag{3.28}$$

*making property **(P1)** easier to verify. (We again wrote S for $r(S')$.)*

We will sometimes define a reversible transformation by G', R and I sets only, in which case we take r to be the *standard* r defined by

$$r(S') = \begin{cases} S' \cup I & \text{if } S' \cap R \neq \emptyset, \\ S' & \text{else.} \end{cases} \tag{3.29}$$

The standard r clearly satisfies properties **(P2)** and **(P3)**, so it will remain to check that properties **(P1)** and **(P4)** hold. In the case that $R \subset I$, or that R is a singleton set, Remark 3.3.6 implies that the second inequality in eq. (3.28) holds too, so it will remain to check the first inequality in order to verify property **(P1)**.

3.3.3 Independence and reducibility

In practice, we will apply a sequence of reversible transformations from G eventually to G'. Say the reversible transformations are

$$G = G_1 \to G_2 \to \ldots \to G_{t+1} = G',$$

with the transformation from G_i to G_{i+1} described by the triplet (I_i, R_i, r_i) for $1 \leq i \leq t$.

If we are given a helpful set in $S' \subset G'$, then

$$S = (r_1 \circ r_2 \circ \ldots \circ r_t)(S')$$

is a helpful set in G by definition of r. In practice the sets I_i will be small (usually $\mathcal{O}(1)$, but at most $\mathcal{O}(\ln n)$), however t may be large. This means that our set S may be large, as the only bound we have is that

$$|S| \leq |S'| + \sum_{i=1}^{t} |I_i|. \tag{3.30}$$

The bound from eq. (3.30) is insufficient for our purposes; in the following we show that under certain conditions the size of S is at most a multiplicative factor (depending on $\max_i |I_i|$) more than S'.

Definition 3.3.7 (Reducibility)**.** Let G and G' be d-regular graphs on the same vertex set, and (V_0, V_1) be a cut. We say that G is *k-reducible* to G' if for all helpful sets $S' \subset V_0$ in G' there exists a helpful set $S \subset V_0$ in G with $|S| \leq k|S'|$ and $h_G(S) \geq h_{G'}(S')$.

Remark 3.3.8 (Multiplicity)**.** *Reducibility is multiplicative in the following sense. If G_1 is k_1-reducible to G_2, and G_2 is k_2-reducible to G_3, then G_1 is $(k_1 k_2)$-reducible to G_3. Similarly for a sequence of t reductions.*

Definition 3.3.9 (Independence of transformations)**.** The sequence of transformations characterised by $(I_i, R_i, r_i)_{i=1}^{t}$ is independent if

$$I_i \cap R_j = \varnothing \tag{3.31}$$

holds for all $i > j$.

The following lemma establishes conditions on a sequence of reversible transformations to obtain a reduction with a small constant, providing us with a more useful control on $|S|$ than eq. (3.30).

Lemma 3.3.10. *Let $q \in \mathbb{Z}^+$ be a positive integer, and consider a sequence of reversible transformations from G_1 eventually to G_{t+1}:*

$$G_1 \to G_2 \to \ldots \to G_{t+1}.$$

Let the transformations be described by the triplets $(I_i, R_i, r_i)_{i=1}^{t}$, satisfying the following:

(C1) *the transformations are independent, i.e. $I_i \cap R_j = \varnothing$ for all $i > j$,*

(C2) *each $v \in G$ is in at most q of the sets R_i.*

Let $k := 1 + \max_{i=1\ldots t} |I_i|$. Then G_1 is (qk)-reducible to G_{t+1}.

Proof. Write $h_i(\cdot)$ for the helpfulness in G_i. Given a set $S_{t+1} \subset G_{t+1}$, for $j = 1 \dots t$ let

$$S_j = (r_j \circ r_{j+1} \circ \dots \circ r_t)(S_{t+1}), \tag{3.32}$$

so that $S_j = r_j(S_{j+1})$. By definition of r (property **(P1)** in Definition 3.3.5) we have that $h_i(S_i) \geq h_{t+1}(S_{t+1})$ for each i, so in particular $h_1(S_1) \geq h_{t+1}(S_{t+1})$. It now remains to show that $|S_1| \leq qk|S_{t+1}|$.

The crucial element here is condition **(C1)** of the lemma. Suppose that for some j we have

$$R_j \cap S_{t+1} = \emptyset.$$

As for each $i > j$ we have $I_i \cap R_j = \emptyset$ (by independence of transformations, **(C1)**), we have

$$R_j \cap S_{j+1} = \emptyset$$

by property **(P2)**, and thus

$$S_j = S_{j+1}$$

by property **(P3)**.

This, in combination with condition **(C2)** implies that at most $q|S_{t+1}|$ of the r_i enlarge their arguments and add vertices to S_1. Each of them adds at most $k - 1$ vertices by definition of k and property **(P2)**, implying that

$$|S_1| \leq qk|S_{t+1}|.$$

Thus G_1 is qk-reducible to G_{t+1}. □

3.4 The $d = 3$ case

In this section we show Theorem 3.1.1 for $d = 3$. Let G be a 3-regular graph, and (V_0, V_1) be a cut of G. Let $\mathcal{A} \subset V_0$ denote the set of vertices that have exactly one cut edge. The following form helpful sets.

(H1) A vertex v in V_0 with more than 1 cut edge, then $\{v\}$ is a helpful set, as in Example 3.3.3.

(H2) A path of three connected $\mathcal{A}$ vertices forms a helpful set of size 3.

(H3) A vertex connected to three $\mathcal{A}$ vertices and one of the $\mathcal{A}$ neighbours connected to another $\mathcal{A}$ vertex forms a helpful set of size 5.

Lemma 3.4.1. *Let G be a 3-regular graph with cut (V_0, V_1). Then either G is $\mathcal{O}(1)$-reducible to a graph that has no connected $\mathcal{A}$ vertices, or we can find a helpful set of size $\mathcal{O}(1)$ in G.*

Proof. If any of the helpful sets **(H1)** to **(H3)** described above exist, we are done. We henceforth assume that such sets are not present. Thus each vertex has 0 or 1 cut edge adjacent to it by **(H1)**.

Consider $x, y \in \mathcal{A}$ connected by an edge. Let u be the remaining neighbour of y in V_0, $u \notin \mathcal{A}$ by **(H2)**. If xu is an edge, we have a helpful set $\{x, y, u\}$ of helpfulness 1 and we are done. Otherwise, by **(H3)** we can find a $v \notin A$ neighbour of u. Consider the transformation of deleting xy and uv edges and adding xu and yv as new edges. See Fig. 3.2.

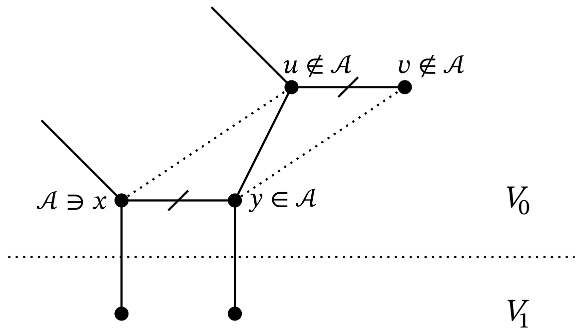


Fig. 3.2 Transforming two connected $\mathcal{A}$ vertices. New edges are shown as dotted, deleted ones are crossed out.

As $u, v \notin \mathcal{A}$, no new $\mathcal{A}\mathcal{A}$ edge is created, and exactly one is removed. We claim the above is a reversible transformation. Write $h(\,.\,)$ and $h'(\,.\,)$ for helpfulness in G and G' respectively. Let S' be a set in G'. If S' contains 0, 1, 3 or 4 of $\{x, y, u, v\}$ then $h(S) = h'(S')$. In the same notation as of Definition 3.3.5 let

$$R = \{x, y\}, \quad I = \{x, y, u\},$$

and

$$r(S') = \begin{cases} S' \cup \{y\} & \text{if } x, u \in S' \\ S' \cup \{x, u\} & \text{if } y, v \in S' \\ S' & \text{else.} \end{cases}$$

It is easy to verify that $h'(S') \leq h'(r(S')) \leq h(r(S'))$ and that this describes a reversible transformation.

Until there is an $\mathcal{A}\mathcal{A}$ edge present, do the transformation described above, letting the final graph be G'. The number of $\mathcal{A}\mathcal{A}$ decreases by 1 after each transformation, so the process terminates or we find a helpful set, which we will show momentarily gives a helpful set of size $\mathcal{O}(1)$ in G. Since initially (in G) we don't have paths on three $\mathcal{A}$ vertices, the transformations are independent.

The conditions of Lemma 3.3.10 with $q = 1$ are satisfied, so the final graph G' has the property that it has no $\mathcal{A}\mathcal{A}$ edges and G is 3-reducible to G'. (By the same argument if we find a helpful set while transforming the graph, the resulting helpful set in G is still of size $\mathcal{O}(1)$.) □

Lemma 3.4.2. *Let G be a 3-regular graph, with cut (V_0, V_1). If*

$$e(V_0, V_1) = (1 + \varepsilon)\frac{1}{3}|V_0|, \tag{3.33}$$

for some $\varepsilon > 0$, then there is a helpful set of size $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ in V_0.

Proof. Apply Lemma 3.4.1. If we obtain a helpful set, we are done. Else relabel G to be G' obtained from the lemma. V_0 does not have adjacent $\mathcal{A}$ vertices, and every vertex is adjacent to at most 1 cut edge, otherwise we have a helpful set.

We construct a new graph H. Form H from $G[V_0]$ by deleting all $\mathcal{A}$ vertices, and place a red edge between the two neighbours in V_0 of each $\mathcal{A}$ vertex deleted (remember, there are no adjacent $\mathcal{A}$ vertices). See Fig. 3.3. So $V(H) = V_0 \setminus \mathcal{A}$.

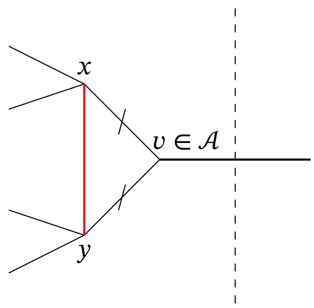


Fig. 3.3 We apply the above transformation for each $\mathcal{A}$ vertex, resulting in graph H.

Crucially, as there are no connected $\mathcal{A}$ vertices in V_0, for each cut edge in G there is a corresponding red edge in H.

The graph H is 3-regular, and has $c := e(V_0, V_1)$ red edges, the rest are black. We have,

$$|H| = |V_0| - c = (2/3 - \varepsilon/3)|V_0|.$$

Hence,

$$c = (1 + \varepsilon)\frac{1}{3}|V_0| = \frac{1/3 + \varepsilon/3}{2/3 - \varepsilon/3}|H| \geq \left(\frac{1}{2} + \frac{\varepsilon}{2}\right)|H|,$$

so by Lemma 3.2.3 there is a subset $S \subset H$ such that graph induced by S has more internal red edges than external black edges in H i.e. $v_H(S) \geq 1$ (as defined in eq. (3.8)), and $|S| = \mathcal{O}(1/\varepsilon)$.

We now show how to construct a set $S' \subset V_0$ with helpfulness $h_G(S') = v_H(S)$. Let $S' \subset V_0$ be the union of S plus all adjacent $\mathcal{A}$ vertices in $G[V_0]$. Each external black edge of S' in H decreases the helpfulness of S' by one; each internal red edge of S' (in H) increases the helpfulness by one; and each external red edge is neutral. Thus

$$h_G(S') = v_H(S) \geq 1,$$

and

$$|S'| \leq 3|S| = \mathcal{O}(1/\varepsilon),$$

so that S' satisfies the lemma, completing the proof.

□

Theorem 3.4.3. *Let G be a 3-regular graph on n vertices. Then,*

$$bw(G) \leq \frac{1}{6}n + \mathcal{O}\left(\sqrt{n}\right).$$

Proof. Let (V_0, V_1) be any bisection of G. We always relabel such that $|V_0| \geq |V_1|$. Let $\varepsilon_0 > 0$ be chosen later. We apply Lemma 3.4.2 to V_0 (larger part) until we can with $\varepsilon \geq \varepsilon_0$, moving the helpful set across to V_1, decreasing the cut size in each step. We relabel if necessary, then repeat. As the cut size decreases in each step, the process terminates. We obtain an (unbalanced) cut (V_0, V_1). By construction,

$$|V_0| = \frac{n}{2} + \mathcal{O}\left(\frac{1}{\varepsilon_0}\right),$$
$$e(V_0, V_1) \leq \frac{1}{3}|V_0|(1+\varepsilon_0) = \frac{1}{6}n + \mathcal{O}\left(\varepsilon_0 n + \frac{1}{\varepsilon_0}\right).$$

Let R be an arbitrary subset $R \subset V_0$ of size $\lceil (|V_0| - |V_1|)/2 \rceil$ to obtain the bisection $(V_0', V_1') = (V_0 \setminus R, V_1 \dot{\cup} R)$. Note that $|R| = \mathcal{O}\left(\frac{1}{\varepsilon_0}\right)$. Thus,

$$e(V_0', V_1') \leq c(V_0, V_1) + 3|R| = \frac{1}{6}n + \mathcal{O}\left(\varepsilon_0 n + \frac{1}{\varepsilon_0}\right). \tag{3.34}$$

The right hand side expression of eq. (3.34) is minimised for $\varepsilon_0 = \frac{1}{\sqrt{n}}$, giving the bound

$$\text{bw}(G) \leq \frac{1}{6}n + \mathcal{O}\left(\sqrt{n}\right).$$

□

3.5 The d even, $d \geq 4$ case

In this section we show Theorem 3.1.1 for $d \geq 4$, d even. For the whole of this section, let

$$d = 2h.$$

For a d-regular graph G equipped with a cut (V_0, V_1), we let $\mathcal{A}, \mathcal{B}, \mathcal{C} \subset V_0$ denote vertices with h, $h-1$, and $\leq h-2$ cut edges (number of edges to V_1) respectively. We shall refer to an '$\mathcal{A}$ vertex' to mean a vertex in the set $\mathcal{A}$; similarly for $\mathcal{B}$ and $\mathcal{C}$.

Lemma 3.5.1. *Let $d \geq 4$ even, G be a d-regular graph. Then, for any (V_0, V_1) bisection of G, there is either a helpful set of size $\mathcal{O}(\ln n)$ in V_0, or there is an injective function $f : \mathcal{A} \to \mathcal{C}$ such that there is a path among $\mathcal{B}$ vertices from each u to $f(u)$. These paths are also vertex-disjoint.*

Proof. Without loss of generality all vertices are incident to at most h cut edges, otherwise they form a helpful set by themselves (Example 3.3.3). Thus all vertices are either $\mathcal{A}$, $\mathcal{B}$ or $\mathcal{C}$ vertices.

The following form helpful sets. We indicate the marginal helpfulness in the figures, as defined in eq. (3.26).

(H1) Two $\mathcal{A}$ vertices joined by a path on $\mathcal{B}$ vertices, including the case of two $\mathcal{A}$ vertices joined by an edge. See Fig. 3.4.

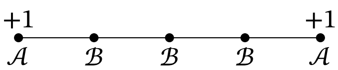


Fig. 3.4 The two end $\mathcal{A}$ vertices contribute at least +1 while the $\mathcal{B}$ vertices are neutral at worst, so the helpfulness is at least 2.

(H2) An $\mathcal{A}$ vertex joined to a cycle of $\mathcal{B}$ vertices on a path of $\mathcal{B}$ vertices, including the case of the path having length 0, so the $\mathcal{A}$ vertex is in the cycle (Fig. 3.5).

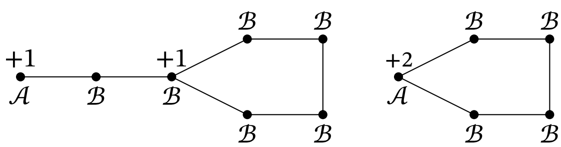


Fig. 3.5 Helpfulness of at least 2.

(H3) A $\mathcal{C}$ vertex joined to $h+1$ $\mathcal{A}$ vertices (Fig. 3.6).

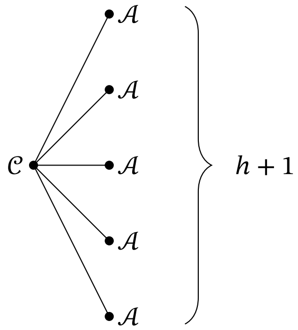


Fig. 3.6 The $\mathcal{A}$ vertices contribute at least $h+1$ in total, while the $\mathcal{C}$ vertex may subtract at most $h-1$, giving a helpfulness of at least 2.

If there are helpful sets of the types above of size at most $2(\log_2 n + 2h) + 1$, we are done. Henceforth we assume no such set exists.

Consider a BFS starting from all the $\mathcal{A}$ vertices (in parallel) traversing along $\mathcal{B}$ vertices only, stopping at depth $\log_2 n + 2h$. The search trees are disjoint, otherwise we can form a helpful set of type **(H1)**; there are no cross-edges in the BFS trees, as otherwise we can form helpful sets of type **(H2)**.

Claim 3.5.2. *All search trees are joined to at least h distinct $\mathcal{C}$ vertices, via paths of length at most* $\log_2 n + 2h$. *(These paths may not be edge disjoint.)*

Proof of claim. Suppose not, so there is a search tree T that is joined to at most $h-1$ distinct $\mathcal{C}$ vertices. The tree has size at least 2, otherwise its a singleton $\mathcal{A}$ vertex which is joined to h $\mathcal{C}$ vertices. Let k be the number of $\mathcal{B}$ vertices with degree 2 in T. The number of edges from T to $\mathcal{C}$ is at least $k(h-1)$, as each degree 2 vertex is joined to $h-1$ of $\mathcal{C}$ vertices. However, it is at most $(h-1)2h$, as each $\mathcal{C}$ vertex has degree $2h$ and T is joined to at most $h-1$ of them. Hence,

$$k(h-1) \le e(T, \mathcal{C}) \le (h-1)2h,$$

and thus,

$$k \le 2h.$$

This implies that all but at most $2h$ levels have vertices with degree 2 (i.e. down degree 1), so vertices in other levels have down-degree ≥ 2 or are leaves. The BFS stops at depth $\log_2 n + 2h$. This means there must be a leaf $v \in T$ that is not at depth $\log_2 n + 2h$, otherwise $|T|$ is larger than n. The only non-$\mathcal{C}$ neighbour of v is its parent. We have that v has $h-1$

cut edges, one edge to its parent, and thus has h distinct $\mathcal{C}$ vertex neighbours, which is a contradiction. *End of proof of claim.* □

Consider a bipartite graph with vertex classes $\mathcal{A}$ and $\mathcal{C}$ and an edge between two vertices if there is a path on $\mathcal{B}$ vertices of length at most $\log_2 n + 2h$ connecting them. By the claim above, each $\mathcal{A}$ vertex has at least h neighbours. By **(H3)**, each $\mathcal{C}$ vertex has at most h neighbours. Hall's marriage theorem [Bol98, Theorem 7, p. 77] says there is an $\mathcal{A}$-saturating matching if for all $X \subset \mathcal{A}$ we have $|\Gamma(X)| \geq |X|$. This is trivially satisfied, as $h|\Gamma(X)| \geq e(X, \Gamma(X)) \geq h|X|$, thus $|\Gamma(X)| \geq |X|$ for all $X \subset \mathcal{A}$.

We let $f : \mathcal{A} \to \mathcal{C}$ be an $\mathcal{A}$-saturating matching in the bipartite graph above, and take the paths between u and $f(u)$ as given by the search trees. The paths are disjoint by construction. □

Remark 3.5.3. *Observe that given a set $S \subset V_0$, $v \in S$ and $b \in \mathcal{B}$ such that vb is an edge, then $S' = S \cup \{b\}$ has $h(S') \geq h(S)$. See Fig. 3.7.*

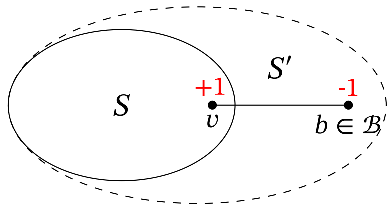


Fig. 3.7 Enlarging S by a $\mathcal{B}$ vertex does not decrease helpfulness. Changes in marginal helpfulness ($h_{S'}(v) - h_S(v)$) indicated. If b has more than one edge to S, the helpfulness actually increases.

Lemma 3.5.4. *Let $d \geq 4$ even, G be a d-regular graph and (V_0, V_1) a bisection of G. Given a matching $f : \mathcal{A} \to \mathcal{C}$ as in Lemma 3.5.1, there exists a d-regular graph G' that is $\mathcal{O}(\ln n)$-reducible to G and has no $\mathcal{A}$ vertices.*

Proof. We have an injection $f : \mathcal{A} \to \mathcal{C}$, such that for each $u \in \mathcal{A}$ there are disjoint paths P_u from u to $f(u)$ on $\mathcal{B}$ vertices.

Consider the path P_u for each $u \in A$. Let $v = f(u)$. We can find vertices $x \in V_1$, $y \in V_0$ such that ux, vy are edges and uy, xv are non-edges. (This holds as u has more neighbours in V_1 than v, and v has more neighbours in V_0 than u, since u is an $\mathcal{A}$ vertex and v is a $\mathcal{C}$ vertex.) See Fig. 3.8.

Consider the transformation on G that replaces the edges ux, vy (shown crossed out) with uv, vx (shown dotted). The new graph is d-regular and has exactly one fewer $\mathcal{A}$ vertex than before, as u becomes a $\mathcal{B}$ vertex and v can become a $\mathcal{B}$ vertex at most.

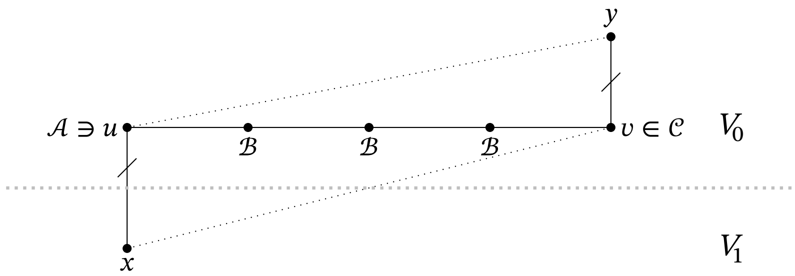


Fig. 3.8 Cartoon showing the transformation of an $\mathcal{A}\mathcal{B}\ldots\mathcal{B}\mathcal{C}$ path.

We show that this a reversible transformation. Let $R = \{v\}$ and I be the vertices of path from u to v. We take the standard r (eq. (3.29)), i.e.

$$r(S') := \begin{cases} S' \cup I & \text{if } v \in S', \\ S' & \text{else.} \end{cases} \tag{3.35}$$

We claim that

$$h'(S') \le h'(r(S')) \le h(r(S')).$$

Indeed, in the case $v \in S'$, the first inequality can been seen by Remark 3.5.3, the second by Remark 3.3.6. Otherwise, if $v \notin S'$, the first inequality is actually an equality, and the second can be seen by inspection.

In any order iterate through all u–$f(u)$ paths. For each path find vertices x and y as above, and apply the transformation as above.

By construction, the sets R and I of different transformations are disjoint, and so the transformations are independent (Definition 3.3.9). Thus, by Lemma 3.3.10 with $q = 1$, we obtain a graph G' that is $\mathcal{O}(\log n)$ reduced from G and has no $\mathcal{A}$ vertices.

□

Lemma 3.5.5. *Let G be a d-regular graph on n vertices, $d \ge 4$ even, and a (V_0, V_1) be a cut of G. Assume that the cut size satisfies*

$$e(V_0, V_1) > \frac{1}{2}|V_0|\left(d - 3 + \frac{3}{d+1} + \varepsilon\right). \tag{3.36}$$

for some $\varepsilon > 0$. Then, we can find a helpful set in V_0 of size $\mathcal{O}\left(\frac{\ln^2 n}{\varepsilon}\right)$ uniformly in ε as $n \to \infty$.

Proof. We may assume ε is sufficiently small. Apply Lemma 3.5.1. If we obtain a helpful set of size $\mathcal{O}(\ln n)$ we are done. Else, we obtain a matching $f : \mathcal{A} \to \mathcal{C}$ that we apply to Lemma 3.5.4 to obtain a graph G' that is $\mathcal{O}(\ln n)$-reducible to G and has no $\mathcal{A}$ vertices. It is thus sufficient to find a helpful set of size $\mathcal{O}\left(\frac{\ln n}{\varepsilon}\right)$ in G' to prove the lemma.

For $0 \leq i \leq d$ let $X_i \subset V_0$ be the vertex induced graph in G that are incident to i cut edges. So X_i is empty for $i \geq h$ (due to the transformations above). Note that the degree of a vertex $v \in X_{h-i}$ is $h + i$.

X_{h-1} is the subgraph induced by the $\mathcal{B}$ vertices. A set $S \subset X_{h-1}$ is helpful if and only if it has $|S| + 1$ or more internal edges. Consider the case that the following holds,

$$|E(X_{h-1})| \geq |X_{h-1}| \left(1 + (h-1)\frac{\varepsilon}{2}\right) \tag{3.37}$$

then by Lemma 3.2.4, with $G = G[X_{h-1}]$, $\Delta = h + 1$, $\beta = \varepsilon/2$, there is a helpful set of size $\mathcal{O}\left(\frac{\ln n}{\varepsilon}\right)$ in G and we are done.

We assume the converse of (3.37) holds, or equivalently, the average number of edges from a $\mathcal{B}$ vertex to $\mathcal{C}$ is at least $(h+1) - 2\left(1 + (h-1)\frac{\varepsilon}{2}\right) = (h-1)(1-\varepsilon)$. Thus,

$$|X_{h-1}|(h-1)(1-\varepsilon) \leq e(\mathcal{B}, \mathcal{C}) \leq \sum_{i=2}^{h-1} |X_{h-i}|(h+i) \tag{3.38}$$

We show that eq. (3.38) gives an upper bound on the cut size that contradicts eq. (3.36) in the premise of the lemma. To this end, we reformulate as linear program with the cut size as the objective function, subject to eq. (3.38) and $\sum_{i=0}^{h} |X_i| = |V_0|$, so that the optimal value of the LP is an upper bound on the cut size. We write the LP in terms of the variables $x_i = \frac{|X_i|}{|V_0|}$ relaxed to $x_i \geq 0$. The linear program is given in Fig. 3.9 and its dual in Fig. 3.10.

$$\begin{aligned}
\text{maximise: } & \sum_{i=1}^{h} x_{h-i}(h-i) \\
\text{subject to: } & \sum_{i=0}^{h-1} x_i = 1 \\
& x_{h-1}(h-1)(1-\varepsilon) \leq \sum_{i=2}^{h} x_{h-i}(h+i) \\
& x_i \geq 0, \quad i = 0, \dots, h-1.
\end{aligned}$$

Fig. 3.9 Linear program relaxation subject to eq. (3.38). The objective function is maximising the cut size.

$$\begin{aligned} &\text{minimise: } \lambda_1 \\ &\text{subject to: } h-1-\lambda_1-\lambda_2(h-1)(1-\varepsilon) \le 0 && (3.39)\\ &\qquad h-i-\lambda_1+\lambda_2(h+i) \le 0, \quad \text{for } i=2,\dots,h && (3.40)\\ &\qquad \lambda_2 \ge 0. \end{aligned}$$

Fig. 3.10 Dual program to Fig. 3.9

Any feasible solution to the dual gives an upper bound on the optimal value of the LP. Let

$$\lambda_1 := h-1-\frac{(h-1)(1-\varepsilon)}{2h+1-\varepsilon(h-1)}, \tag{3.41}$$

$$\lambda_2 := \frac{1}{2h+1-\varepsilon(h-1)}. \tag{3.42}$$

We claim this gives a feasible solution to the dual. It is clear that eq. (3.39) holds with equality. We now check eq. (3.40) for $i=2$. As eq. (3.39) holds with equality, subtracting it from both sides gives that it is equivalent to check that,

$$\lambda_2(h+2+(h-1)(1-\varepsilon)) \le 1,$$

which actually holds with equality. The LHS of eq. (3.40) decreases by $1-\lambda_2 > 0$ as i increases by 1, so it holds for all $i \ge 2$. Thus the solution given by eqs. (3.39) and (3.40) is feasible.

As $(h-1)/(2h+1) < 1/2$, we have that

$$\begin{aligned} \lambda_1 &< h-1-\frac{h-1}{2h+1}+\varepsilon/2 \\ &= h-3/2+\frac{3/2}{2h+1}+\varepsilon/2 \\ &= \frac{1}{2}\left(d-3+\frac{3}{d+1}+\varepsilon\right). \end{aligned}$$

As the solution of the primal is at most λ_1, the cut size is at most $\lambda_1|V_0|$, contradicting eq. (3.36).

□

Theorem 3.5.6. *Let G be a d-regular graph on n vertices, $d \ge 4$ even. Then,*

$$bw(G) \le \left(d-3+\frac{3}{d+1}\right)\frac{n}{4}+\mathcal{O}\left(\sqrt{n}\ln n\right).$$

Proof. Let (V_0, V_1) be any bisection of G. We always relabel such that $|V_0| \geq |V_1|$. Let $\varepsilon_0 > 0$ be chosen later. We apply Lemma 3.5.5 to V_0 (larger part) until we can with $\varepsilon \geq \varepsilon_0$, moving the helpful set across to V_1, decreasing the cut size in each step. We relabel if necessary, then repeat. As the cut size decreases in each step, the process terminates. We obtain an (unbalanced) cut (V_0, V_1). By construction,

$$|V_0| = \frac{n}{2} + \mathcal{O}\left(\frac{\ln^2 n}{\varepsilon_0}\right)$$

$$e(V_0, V_1) \leq \frac{1}{2}|V_0|\left(d - 3 + \frac{3}{d+1} + \varepsilon_0\right) = \frac{n}{4}\left(d - 3 + \frac{3}{d+1}\right) + d\mathcal{O}\left(\varepsilon_0 n + \frac{\ln^2 n}{\varepsilon_0}\right).$$

Let R be an arbitrary subset $R \subset V_0$ of size $\lceil (|V_0| - |V_1|)/2 \rceil$ to obtain the balanced bisection $(V_0', V_1') = (V_0 \setminus R, V_1 \dot{\cup} R)$. Note that $|R| = \mathcal{O}\left(\frac{\ln^2 n}{\varepsilon_0}\right)$. Thus,

$$e(V_0', V_1') \leq e(V_0, V_1) + d|R| = \left(d - 3 + \frac{3}{d+1}\right)\frac{n}{4} + d\mathcal{O}\left(\varepsilon_0 n + \frac{\ln^2 n}{\varepsilon_0}\right). \tag{3.43}$$

The right hand side expression of eq. (3.43) is minimised for $\varepsilon_0 = \frac{\ln n}{\sqrt{n}}$, giving the bound

$$\mathrm{bw}(G) \leq e(V_0', V_1') \leq \left(d - 3 + \frac{3}{d+1}\right)\frac{n}{4} + d\mathcal{O}\left(\sqrt{n}\ln n\right).$$

□

3.6 The d odd, $d \geq 5$ case

In this section we show Theorem 3.1.1 for $d \geq 5$, d odd. For the whole of this section, let h be defined by

$$d = 2h + 1.$$

If a vertex $v \in V_0$ has $h+1$ ore more cut edges it is a helpful set by itself. We assume no such vertices exist from now on. We denote vertices with h, $h-1$, and $h-2$ cut edges as $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ vertices respectively. Vertices with $\leq h-2$ cut edges are denoted $\mathcal{D}$, so that $\mathcal{C} \subset \mathcal{D}$.

For a set $S \subset G$ let $\mathcal{A}(S)$ be the set S enlarged with the $\mathcal{A}$ vertices connected to it,

$$\mathcal{A}(S) := S \cup \left(\Gamma(S) \cap \mathcal{A}\right),$$

where $\Gamma(S)$ is the neighbourhood of S.

We further sub-classify $\mathcal{B}$ and $\mathcal{C}$ vertices by the number of $\mathcal{A}$ neighbours they are incident to. We let $\mathcal{B}_i \subset \mathcal{B}$ and $\mathcal{C}_i \subset \mathcal{C}$ be vertices that are incident to exactly i vertices of type $\mathcal{A}$. A $\mathcal{B}_4$ or a $\mathcal{C}_6$ (or higher subscripts) vertex together with their $\mathcal{A}$ neighbours form a helpful set, so we henceforth assume no such vertices exist.

Table 3.1 below lists some properties of vertices based on classification, for later reference.

v	$e(v, V_1)$	$e(v, V_0 \setminus \mathcal{A})$	$e(v, \mathcal{A})$	$h(\mathcal{A}(v))$
$\mathcal{A}$	h	$h+1$		-1
$\mathcal{B}_3$	$h-1$	$h-1$	3	0
$\mathcal{B}_2$	$h-1$	h	2	-1
$\mathcal{B}_1$	$h-1$	$h+1$	1	-2
$\mathcal{B}_0$	$h-1$	$h+2$	0	-3
$\mathcal{C}_3$	$h-2$	h	3	-2
$\mathcal{C}_2$	$h-2$	$h+1$	2	-3
$\mathcal{D}$	$\leq h-2$	$\geq h+1$		

Table 3.1 Properties of v depending on its type. The rightmost column is a lower bound on $h(A(v))$.

Given a set $S \subset V_0$, $v \in S$ with $v \notin \mathcal{A}$, $b \in V_0 \setminus S$ with $b \in \mathcal{B}_1$ such that vb is an edge, then $S' = S \cup \mathcal{A}(b)$ has $h(S') \geq h(S)$. See Fig. 3.11.

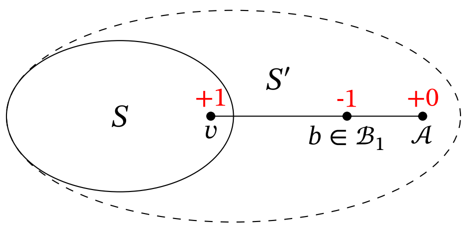


Fig. 3.11 Enlarging S by a $\mathcal{B}_1$ vertex and its $\mathcal{A}$ neighbour does not decrease helpfulness. Changes in marginal helpfulness $(h(v, S) - h(v, S'))$ shown in red.

If b is joined to S by more than one edge, the helpfulness actually increases. If the $\mathcal{A}$ vertex is in S, the helpfulness increases. If v and the $\mathcal{A}$ vertex were to coincide, the helpfulness may actually decrease, but we've specified $v \notin \mathcal{A}$, so this does not happen. We use this observation (and ones similar to it, see Remark 3.6.2) in showing that some transformations are reversible.

3.6.1 Reduction

For a vertex $x \in B_2$ we define $P(x)$ as the number of paths that start from x and terminate in a $\mathcal{D}$ vertex, such that the intermediate vertices are all in $\mathcal{B}_1 \cup \mathcal{B}_2$, and the path has length at most $2\log_2 n$.

Lemma 3.6.1. *There is either a helpful set in G of size $\mathcal{O}(\log n)$ or we can $\mathcal{O}(\log n)$-reduce G to a graph G' such that the following structures are not present in G':*

1. *$\mathcal{A}\mathcal{A}$ edges,*
2. *$\mathcal{B}_2\mathcal{B}_2$ edges,*
3. *$\mathcal{B}_3$ vertices,*
4. *$\mathcal{C}_5$ vertices.*

Further, it holds that,

$$P := \sum_{x \in \mathcal{B}_2} P(x) \geq h|\mathcal{B}_2| + e(\mathcal{B}_2, \mathcal{B}_1). \tag{3.44}$$

Proof. We make a series of reductions for each structure listed, while making sure that we do not recreate a type of structure we have already removed. We deal with each listed forbidden structure in its dedicated subsection.

Transforming $\mathcal{A}\mathcal{A}$ edges

In this subsection we reduce the graph so that it does not contain $\mathcal{A}\mathcal{A}$ edges (i.e. two $\mathcal{A}$ vertices connected by an edge). Let F be the collection of all $\mathcal{A}\mathcal{A}$ edges, represented as an ordered pair. Consider an element $(x, y) \in F$, so that xy is an $\mathcal{A}\mathcal{A}$ edge. Then y has h edges left going to other (non-x) vertices in V_0. Any vertex in V_0 can be connected to at most h edges connecting two $\mathcal{A}$ vertices, otherwise we can find a helpful set of size at most $1 + 2(h + 1)$.

Thus, by Hall's marriage theorem, we can assign a unique neighbour $u \neq x$ of y for all $(x, y) \in F$. Extend the elements of F into triplets by their assigned neighbour u. The elements of F are thus disjoint by construction. If $u \in \mathcal{A} \cup \mathcal{B}_3$, we can form a helpful set of size at most 5, and we are done. By definition $u \notin \mathcal{B}_0$ (as u is attached to y, an $\mathcal{A}$ vertex), thus either $u \in \mathcal{B}_1 \cup \mathcal{B}_2$, or $u \in \mathcal{D}$:

- Case: $u \in \mathcal{B}_1 \cup \mathcal{B}_2$. See Fig. 3.12a. We can find a vertex v that is adjacent to u but not to y, as u has a higher degree in V_0 than y. If $u \in \mathcal{B}_2$ and xu is an edge, then $\{x, y, u\}$ is helpful set of size 3, and we are done. If $u \in \mathcal{B}_1$, then xu is not an edge by definition. Consider the transformation that replaces edges xy and uv with xu and yv, as depicted in Fig. 3.12a. Let $I = R = \{x, y, u\}$ and

$$r(S') = \begin{cases} S' \cup \{y\} & \text{if } x, u \in S', \\ S' \cup \{x, u\} & \text{if } y, v \in S', \\ S' & \text{else.} \end{cases}$$

It is easy to verify that this defines a reversible transformation. (Recall Remark 3.3.6.)

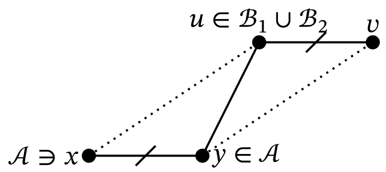
- Case: $u \notin \mathcal{B}$. See Fig. 3.12b. We can find a neighbour v of u that is not adjacent to y, as u has a higher degree in V_0. Similarly, we can find a neighbour $w \in V_1$ of y, that is not adjacent to u. Consider the transformation that replaces edges yw and uv with yv and wu, as depicted in Fig. 3.12b. Let $I = \{x, y\}$, $R = \{u\}$ and

$$r(S') = \begin{cases} S' \cup \{x, y\} & \text{if } u \in S', \\ S' & \text{else.} \end{cases}$$

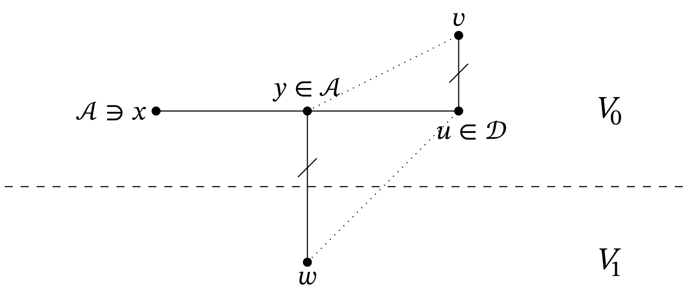
It is straightforward to verify that this defines a reversible transformation.

Observe that both transformations in Fig. 3.12 reduce the number of $\mathcal{AA}$ edges by exactly one. In any order, for each triplet $\{x, y, u\} \in F$, apply the appropriate transformation.

The triples $\{x, y, u\} \in F$ are disjoint by construction, and in both transformations $R, I \subset \{x, y, u\}$, so the transformations are independent (Definition 3.3.9), thus the conditions of Lemma 3.3.10 are satisfied with $q = 1$. Thus, G is $\mathcal{O}(1)$ reducible to the new graph we've just formed, which is without $\mathcal{AA}$ edges.



(a)



(b)

Fig. 3.12 Transforming an $\mathcal{AA}$ edge

Transforming $\mathcal{B}_2\mathcal{B}_2$ edges

We show how to transform the graph to eliminate two $\mathcal{B}_2$ vertices connected by an edge, while not creating a new $\mathcal{A}\mathcal{A}$ edge.

If there is a path on three $\mathcal{B}_2$ vertices, along with their $\mathcal{A}$ neighbours, they form a helpful set of size 9. Assume such paths don't exist, otherwise we are done.

Let xy be an edge with $x, y \in \mathcal{B}_2$. If any of the $\mathcal{A}$ neighbours are shared, we have a helpful set of size at most 5. Let u be any $\mathcal{A}$ neighbour of y. Then u has a neighbour v that is not adjacent to y. (As u has h remaining edges to V_0 and y has $h-1$.) See Fig. 3.13. If $v \in \mathcal{B}_3$, then $\{x, y, u, v\}$ and their $\mathcal{A}$ neighbours form a helpful set. Further, ux is a non-edge, otherwise $\{x, y, u\}$ and their $\mathcal{A}$ neighbours form a helpful set. Consider the transformation as depicted in Fig. 3.13, i.e. replace edges xy and uv with xu and yv. Thus x becomes an $\mathcal{B}_3$ vertex and y remains a $\mathcal{B}_2$ vertex. Note that v can not become an $\mathcal{B}_2$ vertex, thus no new $\mathcal{B}_2\mathcal{B}_2$ edges are created. Also note that no new $\mathcal{A}\mathcal{A}$ edges can be formed.

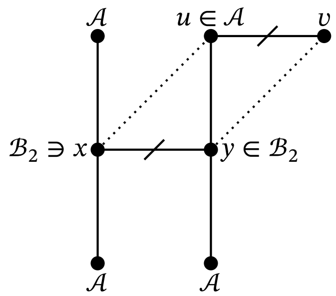


Fig. 3.13 Transformation of a $\mathcal{B}_2\mathcal{B}_2$ edge xy, vertex x becomes $\mathcal{B}_3$ while y stays $\mathcal{B}_2$.

Let $R = \{x, y\}$, $I = \mathcal{A}(x) \cup \mathcal{A}(y)$ and

$$r(S') := \begin{cases} S' \cup \{u\} \cup \mathcal{A}(x) & \text{if } v, y \in S', \\ S' \cup \mathcal{A}(y) & \text{if } u, x \in S', \\ S' & \text{else.} \end{cases}$$

It is straightforward to verify this defines a reversible transformation. In any order, we apply the above transformation to all such $\{x, y\}$ pairs of $\mathcal{B}_2\mathcal{B}_2$ edges. Since initially there are are no $\mathcal{B}_2\mathcal{B}_2\mathcal{B}_2$ paths, the transformations are independent. (As I is a subset of $\mathcal{A}$ vertices, and R is a subset of $\mathcal{B}$ vertices, and these classification do not change.) So the conditions of Lemma 3.3.10 are satisfied, and we obtain a graph without $\mathcal{B}_2\mathcal{B}_2$ and $\mathcal{A}\mathcal{A}$ edges. The reduction is of order $\mathcal{O}(1)$.

Transforming $\mathcal{B}_3$ vertices

In this section we make a $\mathcal{O}(1)$-reduction to a graph that is without $\mathcal{B}_3$ vertices, while not creating $\mathcal{A}\mathcal{A}$ nor $\mathcal{B}_2\mathcal{B}_2$ edges.

Remark 3.6.2. *A $\mathcal{B}_3$ vertex along with its $\mathcal{A}$ neighbours is 0-helpful. If a $\mathcal{B}_3$ vertex is joined to a $\mathcal{B}_2$ vertex, then together (with their $\mathcal{A}$ neighbours) they form a helpful set. However, if one of the $\mathcal{A}$ neighbours of the $\mathcal{B}_3$ vertex is joined to an $\mathcal{B}_2$ vertex, then together (plus $\mathcal{A}$ vertices) they still only form a 0-helpful set.*

In the future, when we say a set forms a helpful set, we implicitly include all their $\mathcal{A}$ neighbours. We will at times omit the calculation of the helpfulness of a set. Usually the easiest way is to look at the marginal contributions of adding each vertex, however we have to be careful: when extending a set, a $\mathcal{B}_1$ vertex when attached via a non-$\mathcal{A}$ vertex does not change helpfulness (as in Fig. 3.11), a $\mathcal{B}_2$ vertex increases it by one. When attached via an $\mathcal{A}$ vertex, a $\mathcal{B}_1$ vertex decreases the helpfulness by one, whereas a $\mathcal{B}_2$ does not change the helpfulness.

A path of type $\mathcal{B}_3\mathcal{A}\mathcal{B}_3$ (along with $\mathcal{A}$ neighbours) is helpful, so assume no such paths exist. Let F be a maximal set of independent $\mathcal{B}_3\mathcal{A}$ edges represented as an ordered pair where the first element is the $\mathcal{B}_3$ vertex. As there are no paths $\mathcal{B}_3\mathcal{A}\mathcal{B}_3$ each $\mathcal{B}_3$ vertex is present exactly once. (Equivalently, we assign an arbitrary $\mathcal{A}$ neighbour to each $\mathcal{B}_3$ vertex.)

Consider an element $(x, a) \in F$, $x \in \mathcal{B}_3$, $a \in \mathcal{A}$ and let $y \neq x$ be a neighbour of vertex a. If $y \in \mathcal{C}_5$, then $\{x, a, y\}$, along with their $\mathcal{A}$ neighbours, forms a helpful set. As there are no $\mathcal{A}\mathcal{A}$ edges, we have that $y \notin \mathcal{A} \cup \mathcal{B}_3 \cup \mathcal{C}_5$.

Consider the case $y \in \mathcal{D}$. We can find a neighbour $w \in V_1$ of a that is not attached to y, as a has more edges to V_1 than y. Similary, we can find a neighbour $v \in V_0$ that is not connected to a. See Fig. 3.14. We claim that transformation that replaces edges vy and aw with av and yw is reversible. Depending on y, we make the following choices with regard to v:

- If $y \in \mathcal{C}_3 \cup \mathcal{C}_4$, we let v be an $\mathcal{A}$ neighbour of y. This is not adjacent to a, as there are no $\mathcal{A}\mathcal{A}$ edges. If $y \in \mathcal{C}_4$, then y does not have a $\mathcal{B}_2$ neighbour, otherwise we can form a helpful set. If $y \in \mathcal{C}_3$ it becomes a $\mathcal{B}_1$, if $y \in \mathcal{C}_4$, it becomes a $\mathcal{B}_2$ after the transformation. In either case, no $\mathcal{B}_2\mathcal{B}_2$ edge can be created.

- Else, we let v be any neighbour of y such that av is not an edge. Such a vertex exists, as y has more neighbours left in V_0 than a. In this case y becomes an $\mathcal{B}_0$ or a $\mathcal{B}_1$ vertex.

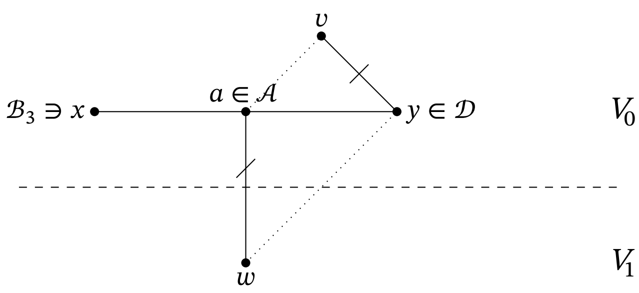


Fig. 3.14 Transforming the $\mathcal{A}$ vertex a to a $\mathcal{B}_0$ (or $\mathcal{B}_1$ if $v \in A$), so that x becomes $\mathcal{B}_2$, whereas y may become a $\mathcal{B}_2$ at most.

In either case no $\mathcal{B}_2\mathcal{B}_2$ nor $\mathcal{A}\mathcal{A}$ edge is created. Let

$$R = \{y\}$$

and

$$I = \mathcal{A}(x)$$

(note that $x, a \in I$), and we take the standard r (eq. (3.29)). It is straightforward to verify this is a reversible transformation.

Iterate over all pairs in F and do as follows. For a pair $\mathcal{B}_3\mathcal{A}$, if the $\mathcal{A}$ vertex is attached to a vertex $y \in \mathcal{D}$, then apply the above transformation and remove the pair from F. If we can't find such a vertex y, move on to the next pair.

In the transformation, the vertex a becomes a $\mathcal{B}$ vertex, whereas the y vertex gains a cut edge, so a vertex may be a "y vertex" at most h times (after that its a helpful set on its own). As R is always a $\mathcal{D}$ vertex, it is trivial that the transformations are independent. Thus, we can apply Lemma 3.3.10 with $q = h$, giving a $\mathcal{O}(1)$ reduction.

For the $\mathcal{B}_3A$ edges that remain in F, it is crucial that now none of them are attached to vertices such as y, i.e. $\mathcal{D}$ vertices. (If the $\mathcal{B}_3\mathcal{A}$ edge is not attached to a $\mathcal{D}$ vertex at time we consider it, it cannot gain a $\mathcal{D}$ neighbour in later transformations.) Thus, in the remaining pairs the $\mathcal{A}$ vertex a only has neighbours in $\mathcal{B}_1$ and $\mathcal{B}_2$.

Consider the case that $y \in \mathcal{B}_1$ and y has no neighbours in $\mathcal{B}_2$. Let v be a neighbour of x that is not connected to y, and u be a non-$\mathcal{A}$ neighbour of y that is not connected to x. See Fig. 3.15. Consider the transformation that replaces xv and yu with vy and xu. Since y has no $\mathcal{B}_2$ neighbours, this does not create a new $\mathcal{B}_2\mathcal{B}_2$ edge.

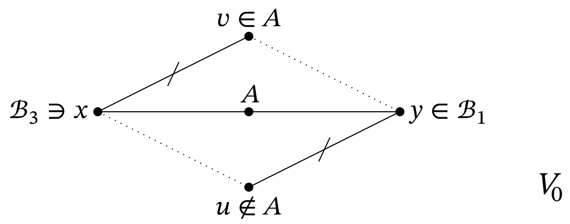


Fig. 3.15 $\mathcal{B}_3$ changes to a $\mathcal{B}_2$, whereas y gains an $\mathcal{A}$ neighbour and becomes $\mathcal{B}_2$. However, no $\mathcal{B}_2\mathcal{B}_2$ edge created as y did not have any $\mathcal{B}_2$ neighbours.

It is straightforward to verity that this is a reversible transformation by

$$R = \{x, y\}, \quad I = \mathcal{A}(x) \cup \mathcal{A}(y),$$

and the standard r (eq. (3.29)). Again, go through all pairs in F and if applicable do the transformation above and delete the pair from F. Again, the transformations are independent, so this is a $\mathcal{O}(1)$ reduction.

Now, paths $\mathcal{B}_3\mathcal{A}\mathcal{D}$, and paths $\mathcal{B}_3\mathcal{A}\mathcal{B}_1$ which are not adjacent to $\mathcal{B}_2$s have been transformed. Thus, all ordered pairs (edges) that remain in F can be continued as a path in one of the two following ways:

- $\mathcal{B}_3\mathcal{A}\mathcal{B}_2$ or,
- $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2$.

Extend each pair in F to a path as above. Note that both these paths are 0-helpful. The paths are vertex disjoint, otherwise we can form a helpful set. Moreover, there is no edge between any of these paths, otherwise we could form a helpful set.

We will apply a transformation to these paths that are very similar to the ones applied to the $\mathcal{B}_3\mathcal{A}$ edges. They are in effect the same with the change that there is a $\mathcal{B}_2$ vertex (or a $\mathcal{B}_1$ and a $\mathcal{B}_2$) in the middle.

Consider a $\mathcal{B}_3\mathcal{A}\mathcal{B}_2$ path and a neighbour of the $\mathcal{B}_2$ vertex, so let $x \in \mathcal{B}_3$, $a \in \mathcal{A}$, $y \in \mathcal{B}_2$ and z a neighbour of y, so that $xayz$ is a path in V_0. Then $z \in \mathcal{D}$ or $z \in \mathcal{B}_0 \cup \mathcal{B}_1$, otherwise we can form a helpful set. If $z \in \mathcal{C}_4 \cup \mathcal{C}_5$, again, we can form a helpful set, so assume otherwise.

Consider the case that $z \in \mathcal{D}$. We can find a neighbour $u \in V_1$ of a that is not attached to z, as a has more edges to V_1 than z. Similarly, we can find a neighbour $v \in V_0$ that is not connected to a. See Fig. 3.16. Consider the transformation that replaces edges aw and zv with av and zw. If $z \in \mathcal{C}_3$, we choose v to be an $\mathcal{A}$ vertex, so that z becomes $\mathcal{B}_1$. This transformation does not create a new $\mathcal{B}_2\mathcal{B}_2$ edge, and removes exactly one $\mathcal{B}_3$ vertex. Let

$$R = \{z\},$$

and

$$I = \mathcal{A}(y) \cup \mathcal{A}(x)$$

note that $a \in I$, and we take the standard (eq. (3.29)) r. It is easy to verify this defines a reversible transformation.

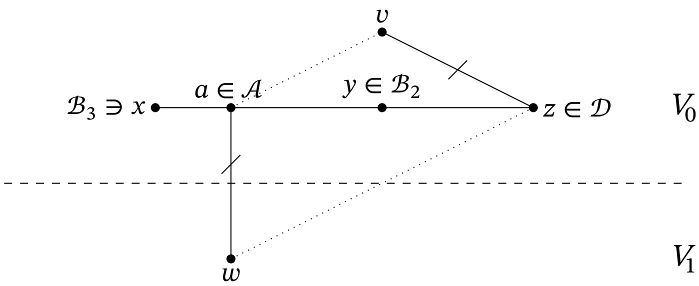


Fig. 3.16 Transforming x into a $\mathcal{B}_2$, as a becomes $\mathcal{B}$, whereas z gains a cut edge.

We can make the same transformation for the paths $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2$ in F, by also adding the $\mathcal{B}_1$ vertex (between a and y) to the set I.

As before, iterate through all paths in F. If we can continue a path with a $z \in \mathcal{D}$ as above, then apply the above transformation. As before, the transformations are independent and so by Lemma 3.3.10 this is a $\mathcal{O}(1)$ reduction. Remove the transformed paths from F.

Now the paths in F (either of type $\mathcal{B}_3\mathcal{A}\mathcal{B}_2$ or $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2$) can only be continued by a vertex $z \in \mathcal{B}_0 \cup \mathcal{B}_1$.

Still considering a $\mathcal{B}_3\mathcal{A}\mathcal{B}_2$ path from F, consider the case $z \in \mathcal{B}_0$. Then, let w be an $\mathcal{A}$ neighbour of x that is not connected to z, and let v be a neighbour of z ($v \notin \mathcal{A}$ as $z \in \mathcal{B}_0$), that is not connected to x. See Fig. 3.17. Consider the transformation that replaces xw and zv with xv and zw. Then x becomes $\mathcal{B}_2$ and z becomes $\mathcal{B}_1$.

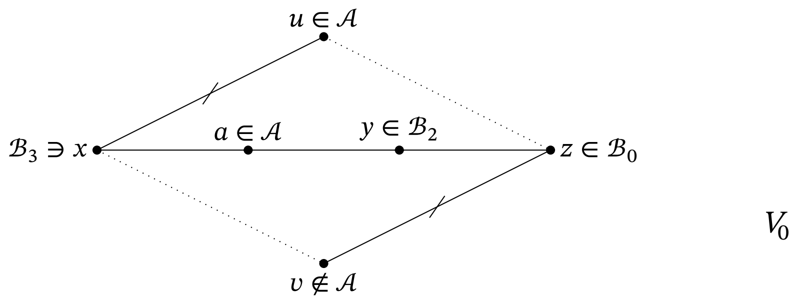


Fig. 3.17 Transforming x into a $\mathcal{B}_2$, while z gains an $\mathcal{A}$ neighbour.

Let

$$R = \{x, z\},$$

$$I = \mathcal{A}(x) \cup \{a\} \cup \mathcal{A}(y) \cup \{z\},$$

and we take the standard r. It is straightforward to verify this gives a reversible transformation. A similar transformation can be defined for paths of type $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2$ in F that transforms the $\mathcal{B}_3$ to a $\mathcal{B}_2$ vertex.

Again, iterate over all paths in F, and if applicable apply the above transformation. The transformations are independent, and by Lemma 3.3.10 this gives a $\mathcal{O}(1)$ reduction. Remove the transformed paths from F.

Now, each element in F can be continued as

- $\mathcal{B}_3\mathcal{A}\mathcal{B}_2\mathcal{B}_1$ or
- $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2\mathcal{B}_1$.

As before, these paths are disjoint, and there are no edges between them (otherwise, we can form a helpful set). Consider the first case. Let xay_1y_2 be a $\mathcal{B}_3\mathcal{A}\mathcal{B}_2\mathcal{B}_1$ path, so $x \in \mathcal{B}_3$, $a \in \mathcal{A}$, $y_1 \in \mathcal{B}_2$ and $y_2 \in \mathcal{B}_1$. Extend to path by a neighbour of y_2, say z. If $z \in \mathcal{B}_2 \cup C_4$ we are done, as the path forms a helpful set.

Consider the case $z \in \mathcal{D}$. See Fig. 3.18. We can find a neighbour $w \in V_1$ of a that is not attached to z, as a has more edges to V_1 than z. Similary, we can find a neighbour $v \in V_0$ that is not connected to a. Consider the transformation that replaces edges aw and vz with av and zw. If $z \in C_3$, then it is not adjacent to a $\mathcal{B}_2$ else we can form a helpful set. If $z \in \mathcal{C}_2$ we let v be an $\mathcal{A}$ neighbour of z, so that no $\mathcal{B}_2\mathcal{B}_2$ edge may be created (as z becomes a $\mathcal{B}_1$). Let

$$R = \{z\}$$

and

$$I = \mathcal{A}(\{y_2, y_1, a\})$$

and we take the standard r. It is straightforward to verify this gives a reversible transformation.

As before, similar transformation for paths $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2\mathcal{B}_1$ that transforms the $\mathcal{B}_3$ to a $\mathcal{B}_2$ vertex can be defined.

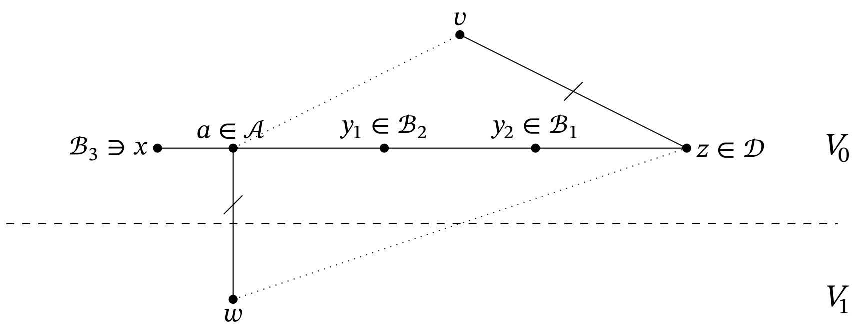


Fig. 3.18 Transforming x into a $\mathcal{B}_2$, as a becomes $\mathcal{B}$, whereas z gains a cut edge.

Again, iterate through all paths in F, and if applicable apply the above transformations. The transformations are independent, thus this is a $\mathcal{O}(1)$ reduction. Remove the transformed paths from F.

It remains to deal with the cases $z \in \mathcal{B}_0$ and $z \in \mathcal{B}_1$. If $z \in \mathcal{B}_1$ then it is not attached to any $\mathcal{B}_2$ vertices, otherwise we can form a helpful set. Consider the following transformation.

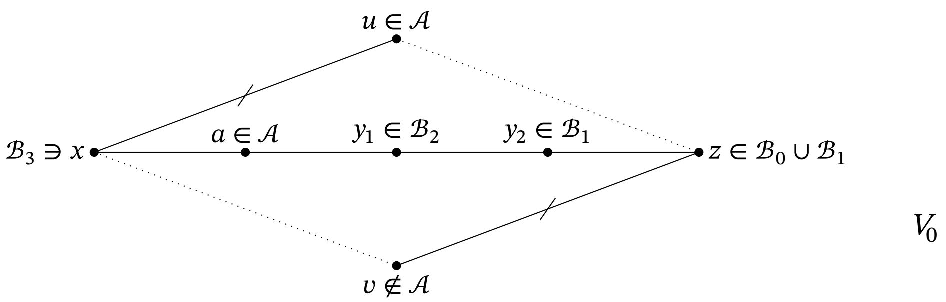


Fig. 3.19 Transforming x into a $\mathcal{B}_2$, while z gains an $\mathcal{A}$ neighbour.

In the case $z \in \mathcal{B}_1$ it was not attached to a $\mathcal{B}_2$ vertex, so no $\mathcal{B}_2\mathcal{B}_2$ may be created. Let

$$R = \{x, z\},$$

$$I = \mathcal{A}(\{x\}) \cup \{a\} \cup \mathcal{A}(\{y_1\}) \cup \mathcal{A}(\{y_2\}) \cup \{z\},$$

and we take the standard r. It is straightforward that this defines a reversible transformation. As before, similar transformation for paths $\mathcal{B}_3\mathcal{A}\mathcal{B}_1\mathcal{B}_2\mathcal{B}_1$ that transforms the $\mathcal{B}_3$ to a

$\mathcal{B}_2$ vertex can be defined. Iterate through all paths in F, and apply the above transformation. As before, this is a $\mathcal{O}(1)$ reduction, and now we have no $\mathcal{B}_3$ vertices left.

Transforming $\mathcal{C}_5$ vertices

The transformation described in Section 3.6.1 holds verbatim if we replace $\mathcal{B}_3$ with $\mathcal{C}_5$. (This holds as a $\mathcal{C}_5$ and a $\mathcal{B}_3$ together with their $\mathcal{A}$ neighbours are both 0-helpful.) For $d = 5$, a $\mathcal{C}_5$ vertex only has $\mathcal{A}$ neighbours. This is the reason we transformed the $\mathcal{B}_3$ vertices "through" their $\mathcal{A}$ neighbours, so that we can apply the same transformations to $\mathcal{C}_5$.

Bound on P

Consider a path $\mathcal{B}_2\mathcal{B}_1^k\mathcal{B}_0$, i.e. a path starting from a $x \in \mathcal{B}_2$ vertex followed by vertices $\{b_i\}_{i=1}^k$ with $b_i \in \mathcal{B}_1$, and terminating in a $y \in \mathcal{B}_0$ vertex. Let $v \in \mathcal{A}$ be an arbitrary $\mathcal{A}$ neighbour of x and u be a neighbour of y that is not joined to x. See Fig. 3.20. We claim that the transformation of replacing the edges xv and yu by xu and yv is a reversible transformation. Let

$$R = \{x, y\}$$

and

$$I = \mathcal{A}(x) \cup \bigcup_{i=1}^{k} \mathcal{A}(b_i) \cup \mathcal{A}(y),$$

and r be the standard choice. We have $|I| = \mathcal{O}(k)$. Observe that this remains a reversible transformation if any of the intermediate vertices b_i are in $\mathcal{B}_2$ instead of $\mathcal{B}_1$. (If at least two of them are in fact in $\mathcal{B}_2$, then could form a helpful set.)

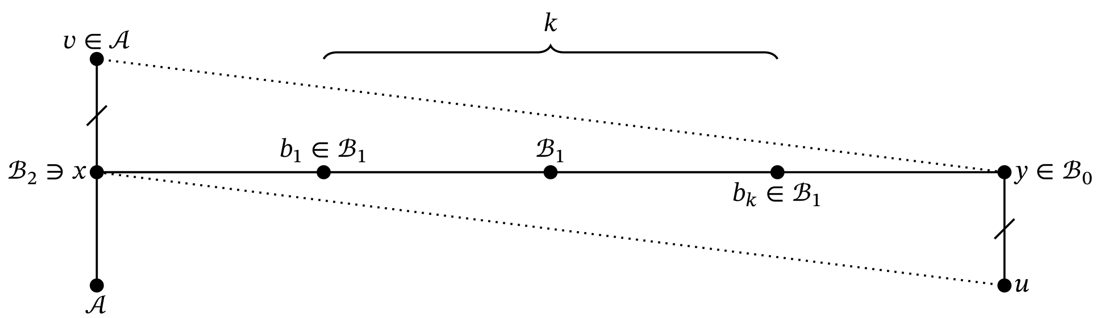


Fig. 3.20 Transforming x and y into $\mathcal{B}_1$ vertices

The transformation changes both x and y into a $\mathcal{B}_1$ vertex, nothing else changes types.

Consider a maximal set of independent $\mathcal{B}_2\mathcal{B}_0$ edges. In any order, we apply the transformation above to them. The transformations are independent, thus this gives a $\mathcal{O}(1)$ reduction. Further, there are no $\mathcal{B}_2\mathcal{B}_0$ edges left. Next, consider a maximal set of $\mathcal{B}_2\mathcal{B}_1\mathcal{B}_0$ paths whose first and last vertices are independent (i.e. the paths may only share intermediate vertices). Again, in any order apply the transformation above to them. This gives a $\mathcal{O}(k)$ reduction, and there are no $\mathcal{B}_2\mathcal{B}_0$ edges or $\mathcal{B}_2\mathcal{B}_1\mathcal{B}_0$ paths left.

Definition 3.6.3. A $\mathcal{B}_2$–$\mathcal{B}_0$ path is a path from $\mathcal{B}_2$ vertex to a $\mathcal{B}_0$ vertex, where all intermediate vertices are $\mathcal{B}_2$ or $\mathcal{B}_1$, of length at most $2\log_2 n$.

Definition 3.6.4. A $\mathcal{B}_2$–$\mathcal{D}$ path is a path from $\mathcal{B}_2$ vertex to a $\mathcal{D}$ vertex, where all intermediate vertices are $\mathcal{B}_2$ or $\mathcal{B}_1$, of length at most $2\log_2 n$.

Recall that for a vertex $x \in \mathcal{B}_2$, $P(x)$ counts the number of $\mathcal{B}_2$–$\mathcal{D}$ paths starting from x. We will transform (reduce) to a graph such that for all $x \in \mathcal{B}_2$,

$$P(x) \geq h + e(v, \mathcal{B}_1). \tag{3.45}$$

and it is then immediate that eq. (3.44) holds.

As there are no $\mathcal{B}_2\mathcal{B}_0$ edges, it is sufficient to show that each $\mathcal{B}_2\mathcal{B}_1$ edge can be continued into two distinct $\mathcal{B}_2$–$\mathcal{D}$ paths. (As a $\mathcal{B}_2$ vertex has h non-$\mathcal{A}$ neighbours, and they all go to a $\mathcal{D}$ or a $\mathcal{B}_1$ vertex, as there are no $\mathcal{B}_2\mathcal{B}_0$ edges.)

Claim 3.6.5. *Let xy be an $\mathcal{B}_2\mathcal{B}_1$ edge, $x \in \mathcal{B}_2$, $y \in \mathcal{B}_0$. Then, xy can be continued to a $\mathcal{B}_2$–$\mathcal{D}$ path in 2 ways, or into as $\mathcal{B}_2$–$\mathcal{B}_0$ in 2 ways ending in distinct $\mathcal{B}_0$s.*

The proof is similar to that of Lemma 3.5.1. This claim would be false had we not removed (transformed) $\mathcal{B}_2\mathcal{B}_1^{\leq 1}\mathcal{B}_0$ paths. If $h = 2$ (i.e. $d = 5$), vertex y has 2 reaming edges to non-$\mathcal{A}$ vertices, so it could be that the graph is locally as depicted in Fig. 3.21. ($\mathcal{A}$ vertices not indicated.)

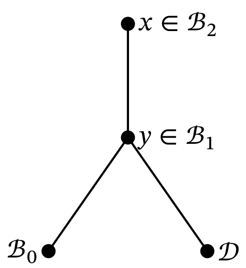


Fig. 3.21 If we allowed $\mathcal{B}_2\mathcal{B}_1\mathcal{B}_0$ paths, the claim would be false.

Proof. Let T be a BFS tree starting from y traversing only on $\mathcal{B}_1$ and $\mathcal{B}_2$ vertices only, excluding x. Continue the tree only up to depth $t := \lceil \log_2 n \rceil + 5$. Note that in $G[V_0 \setminus \mathcal{A}]$ a $\mathcal{B}_1$ vertex has degree $h+1$, and a $\mathcal{B}_2$ vertex has degree h.

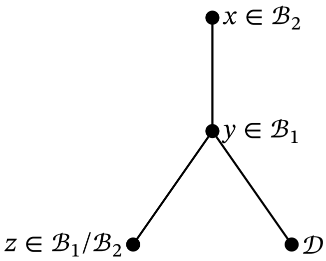
If there is a cross edge in $G[T]$, then we have a helpful set formed by a cycle in T plus a path leading to x from it, and all their $\mathcal{A}$ neighbours. Further, we may only encounter a $\mathcal{B}_2$ vertex in T at most once, otherwise we have three (including x) $\mathcal{B}_2$ vertices connected by $\mathcal{B}_1$ vertices which is a helpful set of size $\mathcal{O}(\log n)$.

A $\mathcal{B}_0$ vertex can have at most 2 edges to T, otherwise we can form a helpful set of size $\mathcal{O}(\log n)$. (If it has 3, take all paths leading back to y, and add the vertex x plus all neighbouring $\mathcal{A}$ vertices.)

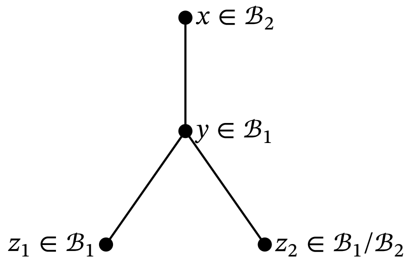
Let k be the number of $\mathcal{B}_1$s in T at depth $< t$ that do not have h children in T. Then they all have at least one edge to $\mathcal{B}_0 \cup \mathcal{D}$. Thus, if $k \geq 4$, then $e(T, \mathcal{B}_0 \cup \mathcal{D}) \geq k \geq 4$, and the claim follows. We henceforth assume $k \leq 3$.

As we stop the BFS at depth t, a leaf of T at depth t may have edges towards $\mathcal{B}_1 \cup \mathcal{B}_2$ vertices in $V_0 \setminus T$. Call a leaf in T a *proper leaf* if it is not at depth t. A proper leaf of type $\mathcal{B}_1$ has $h \geq 2$ edges towards $\mathcal{B}_0 \cup D$, and a proper leaf of type $\mathcal{B}_2$ has $h - 1 \geq 1$ edges towards $\mathcal{D}$. (Remember, there are no $\mathcal{B}_2\mathcal{B}_0$ edges.)

Vertex y has $h \geq 2$ edges remaining to V_0, excluding x and $\mathcal{A}$ vertices. At most one of these is a $\mathcal{D}$ vertex, otherwise we are done. Suppose y has exactly one $\mathcal{D}$ neighbour. As there are no $\mathcal{B}_2\mathcal{B}_1\mathcal{B}_0$ paths, it must have a $\mathcal{B}_1$ or $\mathcal{B}_2$ neighbour as well. Let this $\mathcal{B}$ vertex be z. See Fig. 3.22a.



(a) Exactly one $\mathcal{D}$ neighbours.



(b) No $\mathcal{D}$ neighbours

Fig. 3.22 Vertex y is connected to either 1 or 0 $\mathcal{D}$ vertices

The subtree of T below z does not have a $\mathcal{D}$ neighbour, otherwise we are done (as y already has one $\mathcal{D}$ neighbour). However, as $k \leq 3$, we must encounter a proper leaf of T, otherwise $|T|$ is larger than n. This proper leaf will have $h \geq 2$ distinct $\mathcal{B}_0$ neighbours.

Consider the case that y is adjacent to no $\mathcal{D}$ vertices. Then it has at least one $\mathcal{B}_1$ neighbour, say z_1, and another in $\mathcal{B}_1 \cup \mathcal{B}_2$, say z_2. See Fig. 3.22b. Then, as before we can find a proper leaf in both subtrees rooted at z_1 and z_2. At most one of those proper leaves can be adjacent to $\mathcal{D}$ vertices, or else we are done. Thus at least one the proper leaves has edges to $\mathcal{B}_0$s, thus it must be an $\mathcal{B}_1$ vertex adjacent to $h \geq 2$ vertices of type $\mathcal{B}_0$. □

Let F be the set of $\mathcal{B}_2$ vertices for which eq. (3.45) does not hold. Then, each $v \in F$ must be joined to a $\mathcal{B}_1$ vertex which leads to at most one $\mathcal{D}$ vertex via a $\mathcal{B}_2$–$\mathcal{D}$ as v is connected to two $\mathcal{A}$ vertices, and $\mathcal{B}_1$ and $\mathcal{D}$ vertices only. So by Claim 3.6.5 it leads to two distinct $\mathcal{B}_0$ via a $\mathcal{B}_2$–$\mathcal{B}_0$ path.

Consider a bipartite graph with vertex sets F and $\mathcal{B}_0$, with an edge between two vertices if there exists a $\mathcal{B}_2$–$\mathcal{B}_0$ path between them. By the paragraph above, the left-degree is at least two, the right degree is at most 2 (as each $\mathcal{B}_0$ is the endpoint of at most two $\mathcal{B}_2$–$\mathcal{B}_0$ paths, otherwise can form a helpful set).

By Hall's marriage theorem, there is an F-saturating matching in this bipartite graph. Thus, we can assign a $\mathcal{B}_2$–$\mathcal{B}_0$ path to each $v \in \mathcal{B}_2$ starting from v ending in a unique $\mathcal{B}_0$ vertex. The first and last vertices of these $\mathcal{B}_2$–$\mathcal{B}_0$ paths are distinct, but the paths may intersect. However, each path may intersect at most one other path, otherwise we could form a helpful set. Let the set of these paths be K.

Select a maximal set of vertex-disjoint paths from K. Iterate through them (in any order), and apply the transformation depicted in Fig. 3.20 to them, with $k \leq 2\log_2 n$, then delete the path from K. These transformations are independent, and thus this is a $\mathcal{O}(\log n)$ reduction. The transformations changed some $\mathcal{B}_2$s to $\mathcal{B}_1$s and some $\mathcal{B}_0$s to $\mathcal{B}_1$s. Thus all paths remaining in K are still $\mathcal{B}_2$–$\mathcal{D}$ paths, and are now all vertex-disjoint. Thus, we again iterate through all of them, and obtain a new graph again through a $\mathcal{O}(\log n)$ reduction.

F was the set of all $\mathcal{B}_2$s that violated eq. (3.45). They have all been transformed to a $\mathcal{B}_1$, and thus are no longer violating. The transformation could not cause another $\mathcal{B}_2$ to violate eq. (3.45).

Now we have a graph that satisfies the lemma, and the product of all reductions is $\mathcal{O}(\log n)$, as claimed. □

3.6.2 Relaxation to a Linear Program and result

In this subsection we use Lemma 3.5.4 to reformulate as a linear program to prove the following lemma.

Lemma 3.6.6. *Let G be a d-regular graph on n vertices, $d \geq 5$ odd, and a (V_0, V_1) be a cut of G. Assume that the cut size satisfies*

$$e(V_0, V_1) > \frac{1}{2}|V_0|\left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3} + \varepsilon\right). \tag{3.46}$$

for some $\varepsilon > 0$. Then, we can find a helpful set in V_0 of size $\mathcal{O}\left(\frac{\ln^2 n}{\varepsilon}\right)$ uniformly in ε as $n \to \infty$.

Proof. As in the proof of Lemma 3.5.5, we may assume without loss of generality that ε is sufficiently small. Apply Lemma 3.6.1, we either find a helpful set and we are done, or we have a $\mathcal{O}(\log n)$ reduction to a graph G' as specified in the lemma. Relabel G to be G', and we will remember the $\mathcal{O}(\log n)$ increase in the size of the helpful sets.

First, we note that,

$$e(\mathcal{A}, \mathcal{D}) = e(\mathcal{A}, \mathcal{D} \cup \mathcal{B}_0) = (h+1)|\mathcal{A}| - |\mathcal{B}_1| - 2|\mathcal{B}_2|. \tag{3.47}$$

Let $X_{h-i} \subset V_0$ be the set of vertices incident to $h - i$ cut edges. Thus $\mathcal{D} = \bigcup_{i=2}^{h} X_{h-i}$. Consider a vertex $v \in X_{h-i}$. Then v by itself is $-(2i+1)$ helpful. Recall from eq. (3.44), P counts the total number of $\mathcal{B}_2$–$\mathcal{D}$ paths. If there is a $\mathcal{B}_2$–$\mathcal{D}$ ending in v, then extending v by the path increases its helpfulness by at least one. Similarly, extending v by a neighbouring $\mathcal{A}$ vertex also increases its helpfulness by one. So the total number of $\mathcal{A}$ neighbours plus the number of $\mathcal{B}_2$–$\mathcal{D}$ paths ending in v is at most $2i + 1$, otherwise we can form a helpful set of size $\mathcal{O}(\log n)$. Thus,

$$e(\mathcal{A}, \mathcal{D}) + P \leq \sum_{i=2}^{h} (2i+1)|X_{h-i}|. \tag{3.48}$$

From eqs. (3.44), (3.47) and (3.48), and $h \geq 2$, it follows that,

$$(h+1)|\mathcal{A}| - |\mathcal{B}_1| + e(\mathcal{B}_2, \mathcal{B}_1) \leq \sum_{i=2}^{h} (2i+1)|X_{h-i}|. \tag{3.49}$$

Further, as there are no $\mathcal{C}_5$ vertices, and a vertex in X_{h-i} has at most $2i + 1$ edges to $\mathcal{A}$ vertices otherwise we can form a helpful set, we have that,

$$(h+1)|\mathcal{A}| - |\mathcal{B}_1| - 2|\mathcal{B}_2| = e(\mathcal{A}, \mathcal{D}) \leq 4|X_{h-2}| + \sum_{i=3}^{h} (2i+1)|X_{h-i}|. \tag{3.50}$$

A set $S \subset \mathcal{B}_1$ with $|S| + 1$ internal edges is helpful (along with its $\mathcal{A}$ neighbours). Thus the number of edges within $\mathcal{B}_1$ vertices is at most $(1 + \varepsilon)|\mathcal{B}_1|$, otherwise we have a helpful set of size $\mathcal{O}(\log n/\varepsilon)$ by Lemma 3.2.4. Thus a $\mathcal{B}_1$ vertex has at least $(h-1-\varepsilon)$ edges towards non-$\mathcal{A}$ and non-$\mathcal{B}_1$ vertices. Hence,

$$e(\mathcal{B}_1, \mathcal{B}_0 \cup \mathcal{D}) \geq (h - 1 - \varepsilon)|\mathcal{B}_1| - e(\mathcal{B}_1, \mathcal{B}_2). \tag{3.51}$$

Further, as each $\mathcal{B}_2$ vertex has 2 edges towards $\mathcal{A}$, and h towards $V_0 \setminus \mathcal{A}$, we have that,

$$e(\mathcal{B}_2, \mathcal{B}_0 \cup \mathcal{D}) \geq h|\mathcal{B}_2| - e(\mathcal{B}_1, \mathcal{B}_2). \tag{3.52}$$

Remember that $\mathcal{D} = \bigcup_{i=2}^{h} X_{h-i}$. Thus, as $h \geq 2$, from eqs. (3.47), (3.51) and (3.52) we have

$$(h + 1)|\mathcal{A}| + (h - 2 - \varepsilon)|\mathcal{B}_1| - 2e(\mathcal{B}_1, \mathcal{B}_2) \leq e(A \cup \mathcal{B}_2 \cup \mathcal{B}_1, \mathcal{B}_0 \cup \bigcup_{i=2}^{h} X_{h-i}) \tag{3.53}$$

$$\leq (h + 2)|\mathcal{B}_0| + \sum_{i=2}^{h}(h + i + 1)|X_{h-i}|, \tag{3.54}$$

where the second inequality holds as a vertex in X_{h-i} has $h + i + 1$ edges towards V_0.

We rescale the variables by $|V_0|$: we let $a = |\mathcal{A}|/|V_0|$, $b_i = |\mathcal{B}_i|/|V_0|$ $x_i = |X_i|/|V_0|$, and $m = e(\mathcal{B}_2, \mathcal{B}_1)/|V_0|$. Then, we reformulate as a linear program Primal(ε) with parameter ε (see Fig. 3.23), whose objective function is maximising the cut size subject to eqs. (3.49), (3.50) and (3.53). We denote the value of the optimal solution to Primal(ε) by opt(ε). Thus opt(ε) gives an upper bound on the cut size (scaled by $|V_0|$), as we are maximising for the cut size, i.e.

$$e(V_0, V_1) \leq |V_0| \cdot \text{opt}(\varepsilon). \tag{3.59}$$

Consider the following solution to Primal(0). Let

$$a = \frac{h + 2}{2h^2 + 4h + 3},$$
$$b_1 = \frac{(h + 1)(h + 2)}{2h^2 + 4h + 3},$$
$$b_0 = \frac{(h + 1)(h - 1)}{2h^2 + 4h + 3},$$

$$
\begin{aligned}
\text{maximise:} \quad & \sum_{i=0}^{h-2} i x_i + (h-1)(b_0+b_1+b_2) + ha \\
\text{subject to:} \quad & \sum_{i=0}^{h-2} x_i + b_0 + b_1 + b_2 + a = 1 && (3.55)\\
& (h+1)a - b_1 - 2b_2 \le 4x_{h-2} + \sum_{i=3}^{h}(2i+1)x_{h-i} && (3.56)\\
& (h+1)a - b_1 + m \le \sum_{i=2}^{h}(2i+1)x_{h-i} && (3.57)\\
& (h+1)a + (h-2-\varepsilon)b_1 - 2m \le (h+2)b_0 + \sum_{i=2}^{h}(h+i+1)x_{h-i} && (3.58)\\
& m, a, b_2, b_1, b_0, x_i \ge 0, \quad i = 0, \dots, h-2.
\end{aligned}
$$

Fig. 3.23 The linear program Primal(ε) with parameter ε.

and the other variables are 0. It is straightforward to verify this solution is feasible. The value of the solution is as follows:

$$
\begin{aligned}
ha + (h-1)(b_1+b_0) &= \frac{h(h+2) + (h-1)(h+1)(2h+1)}{2h^2+4h+3} \\
&= \frac{2h^3+2h^2-1}{2h^2+4h+3} \\
&= h - 1 + \frac{h+2}{2h^2+4h+3} \\
&= \frac{1}{2}\left(d - 3 + \frac{2(d+3)}{d^2+2d+3}\right). && (3.60)
\end{aligned}
$$

The dual program Dual(ε) is given in Fig. 3.24. We have the following correspondence between variables and constraints of the primal and dual:

- λ_1–eq. (3.55); λ_2–(3.56); λ_3–eq. (3.57); λ_4–eq. (3.58),
- m–eq. (3.61); a–eq. (3.62); b_2–eq. (3.63); b_1–eq. (3.64); b_0–eq. (3.65); x_{h-2}–eq. (3.66); x_{h-i}–eq. (3.67).

minimise: λ_1

subject to:

$$-\lambda_3 + 2\lambda_4 \leq 0 \tag{3.61}$$
$$h - \lambda_1 - (h+1)\lambda_2 - (h+1)\lambda_3 - (h+1)\lambda_4 \leq 0 \tag{3.62}$$
$$(h-1) - \lambda_1 + 2\lambda_2 \leq 0 \tag{3.63}$$
$$(h-1) - \lambda_1 + \lambda_2 + \lambda_3 - (h-2-\varepsilon)\lambda_4 \leq 0 \tag{3.64}$$
$$(h-1) - \lambda_1 + (h+2)\lambda_3 \leq 0 \tag{3.65}$$
$$(h-2) - \lambda_1 + 4\lambda_2 + 5\lambda_3 + (h+3)\lambda_4 \leq 0 \tag{3.66}$$
$$(h-i) - \lambda_1 + (2i+1)\lambda_2 + (2i+1)\lambda_3 + \lambda_4(h+i+1) \leq 0 \quad i = 3, \dots, h. \tag{3.67}$$
$$\lambda_2, \lambda_3, \lambda_4 \geq 0$$

Fig. 3.24 The dual linear program Dual(ε) to Primal(ε).

We claim the following is a feasible solution to Dual(0). Observe from eq. (3.60) that λ_1 is equal to the solution of the primal. Let

$$\lambda_1 = h - 1 + \frac{h+2}{2h^2+4h+3},$$
$$\lambda_2 = \frac{h+2}{2(2h^2+4h+3)},$$
$$\lambda_3 = \frac{3h-2}{2(2h^2+4h+3)},$$
$$\lambda_4 = \frac{1}{2h^2+4h+3}.$$

It is straightforward to verity that eqs. (3.61) to (3.66) are satisfied. Consider eq. (3.67) with $i = 3$. Multiplying both sides by $2h^2 + 4h + 3$ (the RHS remains 0), the LHS evaluates to

$$-(2h^2 + 4h + 3) - (h+2) + 10h + (h+3) = -4h^2 + 6h + 2, \tag{3.68}$$

which is negative for $h \geq 2$, i.e. for $d \geq 5$, as required. The $i = 3$ case implies that the equation holds for all $i \geq 3$, as the change in the LHS for an increase in i by 1 (or equivalently the derivative with respect to i) is

$$-1 + 2\lambda_2 + 2\lambda_3 + \lambda_4 = -\frac{2h^2+2}{2h^2+4h+3} \leq 0.$$

By strong duality, opt(ε) is also the solution of Dual(ε), so that opt(0) $= \lambda_1$. Note that ε only appears in eq. (3.64) in the constraints of Dual(ε). Thus given the solution by λ for

Dual(0), it is easy to see that $\lambda'_1 = \lambda_1 + \varepsilon\lambda_4$, and $\lambda'_i = \lambda_i$ for $i \geq 2$ is a feasible solution for Dual(ε). Thus

$$\mathrm{opt}(\varepsilon) \leq \lambda'_1 = \lambda_1 + \varepsilon\lambda_4 < \lambda_1 + \varepsilon/2.$$

Remark 3.6.7. *By standard theory on sensitivity analysis (see e.g. [BT97, Chapter 6, section "Change in the Constraint Matrix"]), we actually have* $\frac{\partial}{\partial\varepsilon}\,\mathrm{opt}(\varepsilon)\Big|_{\varepsilon=0} = b_1 \cdot \lambda_4$.

Hence, by eq. (3.59),

$$e(V_0, V_1) \leq \mathrm{opt}(\varepsilon) < V_0(\lambda_1 + \varepsilon/2) = \frac{1}{2}|V_0|\left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3} + \varepsilon\right), \tag{3.69}$$

contradicting eq. (3.46).

□

Theorem 3.6.8. *Let G be a d-regular graph on n vertices, $d \geq 5$ odd. Then,*

$$bw(G) \leq \left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3}\right)\frac{n}{4} + \mathcal{O}\left(\sqrt{n}\ln n\right).$$

Proof. Let (V_0, V_1) be any bisection of G. We always relabel such that $|V_0| \geq |V_1|$. Let $\varepsilon_0 > 0$ be chosen later. We apply Lemma 3.6.6 to V_0 (larger part) until we can with $\varepsilon \geq \varepsilon_0$, moving the helpful set across to V_1, decreasing the cut size in each step. We relabel if necessary, then repeat. As the cut size decreases in each step, the process terminates. We obtain an (unbalanced) cut (V_0, V_1). By construction,

$$|V_0| = \frac{n}{2} + \mathcal{O}\left(\frac{\ln^2 n}{\varepsilon_0}\right)$$

$$e(V_0, V_1) \leq \frac{1}{2}|V_0|\left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3} + \varepsilon_0\right) = \frac{n}{4}\left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3}\right) + d\mathcal{O}\left(\varepsilon_0 n + \frac{\ln^2 n}{\varepsilon_0}\right).$$

Let R be an arbitrary subset $R \subset V_0$ of size $\lceil(|V_0| - |V_1|)/2\rceil$ to obtain the balanced bisection $(V'_0, V'_1) = (V_0 \setminus R, V_1 \dot\cup R)$. Note that $|R| = \mathcal{O}\left(\frac{\ln^2 n}{\varepsilon_0}\right)$. Thus,

$$e(V'_0, V'_1) \leq e(V_0, V_1) + d|R| = \left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3}\right)\frac{n}{4} + d\mathcal{O}\left(\varepsilon_0 n + \frac{\ln^2 n}{\varepsilon_0}\right). \tag{3.70}$$

The right hand side expression of eq. (3.70) is minimised for $\varepsilon_0 = \frac{\ln n}{\sqrt{n}}$, giving the bound

$$\mathrm{bw}(G) \leq e(V'_0, V'_1) \leq \left(d - 3 + \frac{2(d+3)}{d^2 + 2d + 3}\right)\frac{n}{4} + d\mathcal{O}\left(\sqrt{n}\ln n\right).$$

□

3.7 Acknowledgements

We have made use of the *JuMP* modelling language [DHL17] embedded in the *Julia* programming language [BEKS17].

Chapter 4

Maximum cut of random d-regular graphs

4.1 Introduction

Given a graph $G = (V, E)$, a *cut* is a partition of the vertex set V into two disjoint parts (V_0, V_1). A *bisection* is a cut where the two parts are as equal as possible, i.e. if $|V_0|$ and $|V_1|$ differ by at most 1. The *cut size* of a cut is the number of edges $e(V_0, V_1)$ from V_0 to V_1.

The *maximum cut*, denoted $\mathrm{MaxCut}(G)$, of G is the size of the cut with maximum cut size. The *maximum bisection*, denoted $\mathrm{MaxBis}(G)$, respectively *minimum bisection*, denoted $\mathrm{MinBis}(G)$, is the size of the bisection with maximum, respectively minimum cut size. The minimum bisection is also known as *bisection width* (as we refer to it in Chapter 3). Trivially, $\mathrm{MaxCut}(G) \geq \mathrm{MaxBis}(G)$.

All three problems are hard combinatorial optimization problems; their decision problems are all NP-complete. They remain NP-complete even when restricted to d-regular graphs, $d \geq 3$.

The random d-regular graph $\mathbb{G} = \mathbb{G}(n, d)$ is a graph drawn uniformly at random from the set of all possible d-regular graphs on n vertices (provided that dn is even). In this chapter our focus is on the quantity $\mathrm{MaxCut}(\mathbb{G})$ and we derive a high probability upper bound on the random variable $\mathrm{MaxCut}(\mathbb{G})$.

The precise analysis of the quantities $\mathrm{MaxCut}(\mathbb{G})$, $\mathrm{MaxBis}(\mathbb{G})$, and $\mathrm{MinBis}(\mathbb{G})$ is a long-standing open problem. In the $d \to \infty$ limit it has been proved that the main correction term, P in Theorem 4.1.1 below, arises from the 'Parisi formula' in the Sherrington-Kirkpatrick model [Tal06].

Theorem 4.1.1 ([DMS17, Theorems 1.5 and 1.6]). *There exist a constant $P > 0$, such that for any $d \geq 3$, as $n \to \infty$, we have w.h.p.,*

$$\frac{\mathrm{MaxCut}(\mathbb{G})}{dn/2} = \frac{1}{2} + \frac{P}{\sqrt{d}} + o_d\left(1/\sqrt{d}\right),$$
$$\frac{\mathrm{MaxBis}(\mathbb{G})}{dn/2} = \frac{1}{2} + \frac{P}{\sqrt{d}} + o_d\left(1/\sqrt{d}\right),$$
$$\frac{\mathrm{MinBis}(\mathbb{G})}{dn/2} = \frac{1}{2} - \frac{P}{\sqrt{d}} + o_d\left(1/\sqrt{d}\right).$$

Note that the above theorem has no implications for any fixed d. The constant P is expressed analytically and $P \approx 0.76321$. This result is especially striking, as for a given graph G, there is no obvious combinatorial relation between the bisections maximising and minimising the cut size.

Our contribution is deriving explicit upper bounds on $\mathrm{MaxBis}(\mathbb{G}(n,d))$. The precise statement is in the following section.

4.2 Main result

Our main result will be expressed via a closed formula that arises from analysing a random walk. In the following, let

$$\mathcal{M}_d = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots \\ \frac{1}{2} & 0 & \frac{1}{2} & \ddots & & \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \\ \vdots & & & 0 & 1 & 0 \end{bmatrix}. \tag{4.1}$$

be a band matrix of size $(d+1) \times (d+1)$, so that M_d is a transition matrix. Let t be a dummy variable. We define the matrix

$$\mathcal{A}_{d,\alpha} = (1 - 2\alpha)\mathrm{id} + 2\alpha t \mathcal{M}_d$$

for $0 < \alpha < 1/2$ and id denotes the identity matrix. Further, define vectors ζ and ξ as

$$\zeta = \begin{bmatrix} 1, & 0, & 0, & \cdots & \end{bmatrix} \in \mathbb{R}^{1\times(d+1)}, \tag{4.2}$$

$$\xi = \begin{bmatrix} 1, & t^{-1}, & t^{-2}, & t^{-3}, & \cdots \end{bmatrix}^T \in \mathbb{R}^{(d+1)\times 1}. \tag{4.3}$$

We are now ready to express the function in terms of which our main result is stated. For $\alpha \in (0, 1/2)$ and $z \in (0, 1)$ let

$$F_d(\alpha, z) = -\frac{\log\left(\zeta \mathcal{A}^d_{d,\alpha} \xi|_{t=z^{1/2}}\right)}{\log z} + \frac{d}{2}\frac{\log\left(1 - 2\alpha^2 + 2\alpha^2 z\right)}{\log z}. \tag{4.4}$$

Recall that $\mathbb{G}(n, d)$ is the random d-regular graph. Our contribution is the following theorem.

Theorem 4.2.1. *Let* $\mathrm{MaxCut}(\mathbb{G})$ *be the size of the maximum cut of* $\mathbb{G}(n, d)$*. Then, w.h.p.*

$$\frac{\mathrm{MaxCut}(\mathbb{G})}{dn/2} \leq 1 + \frac{2}{d} \inf_{\substack{\alpha\in(0,1/2)\\ z\in(0,1)}} \left(F_d(\alpha, z)\right) + o(1). \tag{4.5}$$

Note that $|E(\mathbb{G})| = dn/2$, and that the infimum of $F_d(\alpha, z)$ will be negative. We give numerical solutions to the optimization problem eq. (4.5) for $d = 3 \ldots 10$. Due to the complexity of the formula eq. (4.4) — logarithms of polynomials in α and z —- we can only derive numerical results.

Corollary 4.2.2. $\mathrm{MaxCut}(\mathbb{G})/|E(\mathbb{G})|$ *is numerically upper-bounded (w.h.p.) by the following values.*

d	3	4	5	6	7	8	9	10
$\mathrm{MaxCut}(\mathbb{G})/\lvert E\rvert$	0.9241	0.8683	0.8350	0.8049	0.7851	0.7659	0.7523	0.7388

Table 4.1 Upper bounds on the maximum cut size of a random d-regular graph, for $d = 3 \ldots 10$. The values for α, z and $F_d(\alpha, z)$ can be found in Table 4.2 in Section 4.5.

The values in Table 4.1 match the conjectured values (obtained from numerical solutions to the one-step replica breaking equations) from [ZB10]. The bound from $d = 3$ improves on the bound of 0.9351 from [McK82], [Hla06] and 0.9319 from [Sor20].

4.3 Statistical physics formulation

This section (Section 4.3) follows the excellent description of the interpolation method from [ACG19].

Note on notation: we will write for example $\prod 1-x$ to mean $\prod(1-x)$, to declutter the layout; the meaning should always be clear.

The general approach we follow is by now standard in a combinatorial literature inspired by statistical physics, but less familiar in combinatorics generally. Consequently, while where possible we will refer to other works for lengthy details. Nonetheless, we provide an introduction to statistical physics and give an outline of the interpolation argument.

Ayre, Coja-Oghlan, and Greenhill [ACG19] used an adaptation of the interpolation method by [SSZ16] to lower-bound the chromatic number of random graphs. Our interpolation argument is a special case of theirs and so we will point the reader to the relevant sections of [ACG19], keeping the same notation.

Given a multigraph $G = (V, E)$, a mapping $\sigma \in \{\pm 1\}^V$ of each vertex to $+1$ or -1 (called *spins*) is called a *configuration*. A configuration σ naturally gives rise to a cut of G, by letting the two vertex classes be the vertices to which σ assigns $+1$ and -1 respectively. We let *total energy* (also known as the *Hamiltonian*) of a configuration σ be

$$\mathcal{H}_G(\sigma) = \sum_{uv \in E(G)} \mathbf{1}\{\sigma_u = \sigma_v\}, \tag{4.6}$$

so that $\mathcal{H}_G(\sigma)$ counts the number of uncut (unsatisfied) edges in the cut given by σ.

The *Ising antiferromagnet* with inverse temperature (parameter) $\beta \geq 0$ on G is a probability distribution $\mu_{G,\beta}$ on all configurations $\Omega = \{\pm 1\}^V$ defined by

$$\mu_{G,\beta}(\sigma) = \frac{1}{Z_\beta(G)} \exp(-\beta \mathcal{H}_G(\sigma)), \tag{4.7}$$

$$Z_\beta(G) = \sum_{\sigma \in \{\pm 1\}^V} \exp(-\beta \mathcal{H}_G(\sigma)). \tag{4.8}$$

The function $Z_\beta(G)$ is called the *partition function*. The distribution assigns a penalty term $\exp(-\beta)$ for each unsatisfied (uncut) edge in the configuration σ. At $\beta = 0$, the distribution becomes the uniform distribution on all configurations. In the limit $\beta \to \infty$, the probability becomes concentrated on configurations with minimal energy, i.e. those corresponding to the maximum cut of G. This is the limit we will be interested in. We have that,

$$\frac{2}{dn}\,\mathrm{MaxCut}(G) = 1 - \frac{2}{dn}\min_{\sigma\in\{\pm1\}^V}\mathcal{H}_G(\sigma) \le 1 + \frac{2}{\beta dn}\log Z_\beta(G). \tag{4.9}$$

The last inequality follows from $\log Z_\beta(G) \ge \log(\exp(-\beta\mathcal{H}(\sigma_M))) = -\beta\mathcal{H}_G(\sigma_M)$, where $\sigma_M = \operatorname{argmin}_{\sigma\in\{\pm1\}^V}\mathcal{H}_G(\sigma)$. Our goal is to upper-bound the log-partition function $\log Z_\beta(\mathbb{G})$ to obtain an upper bound on $\mathrm{MaxCut}(\mathbb{G})$ via eq. (4.9). To this end, we will deduce an upper bound on $\mathbb{E}[\log Z_\beta(\mathbb{G})]$ and apply Proposition 4.3.1 below which says that $\log Z_\beta(\mathbb{G})$ is concentrated around its mean.

Proposition 4.3.1 ([ACG19, Proposition 2.1])**.** *For any $\beta,\delta > 0$, there exists $\kappa > 0$ such that,*

$$\mathbb{P}\Big[\big|\log Z_\beta(\mathbb{G}) - \mathbb{E}[\log Z_\beta(\mathbb{G})]\big| > \delta n\Big] \le \exp(-\kappa n), \tag{4.10}$$

for n sufficiently large.

However, analysing the log-partition function directly is very hard due to sheer number of terms. To analyse $\log(Z_\beta(\mathbb{G}))$ we instead work in the following configuration model: fix $\varepsilon > 0$ small, and let

$$\mathbf{m} \sim \mathrm{Po}_{\le dn/2}((1-\varepsilon)dn/2) \tag{4.11}$$

be a Poisson random variable of rate $(1-\varepsilon)dn/2$ conditioned on not exceeding $dn/2$. Define $\mathbf{G} = \mathbf{G}(n,d)$ to be the random multigraph obtained by choosing a matching $\mathbf{L}$ of size $\mathbf{m}$ uniformly from the complete graph on $V_n \times [d]$. Then, for each edge $\{(u,i),(v,j)\} \in \mathbf{L}$ we add edge uv (with multiplicity) to $\mathbf{G}$.

For our purposes — the upper bound that can be derived from eq. (4.9) — the following proposition says that the true d-regular model $\mathbb{G}$ and the configuration model $\mathbf{G}$ are equivalent, as our upper bound will come from an upper bound on $\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\log Z_\beta(\mathbb{G})\right]$.

Proposition 4.3.2 ([ACG19, Corollary 2.2])**.** *For any $\beta > 0$, we have*

$$\lim_{\varepsilon\to0}\lim_{n\to\infty}\frac{1}{n}\Big|\mathbb{E}[\log Z_\beta(\mathbb{G})] - \mathbb{E}[\log Z_\beta(\mathbf{G})]\Big| = 0.$$

Finally, for all $d,\beta > 0$ let

$$\Phi_d(\beta) = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\log Z_\beta(\mathbb{G})\right]. \tag{4.12}$$

See e.g. [Pan13, Theorem 1.1] for a proof that the limit exists. With reference to eq. (4.9), Propositions 4.3.1 and 4.3.2, the key quantity of interest is $\Phi_d(\beta)$.

4.3.1 The interpolation method

Still following [ACG19].

Recall that our aim is to upper-bound $\log(Z_\beta(\mathbf{G}))$. To analyse the log-partition function $\log(Z_\beta(\mathbf{G}))$, we will instead compare it to a graph $\mathbf{G}_1$, where the dependencies (in the partition function) are more manageable. We construct a family of random graphs $(\mathbf{G}_t)_{t\in[0,1]}$ such that we can easily relate (the partition functions of) $\mathbf{G}$ and $\mathbf{G}_0$. For the graphs $\mathbf{G}_t$ we will actually define a generalized partition function, which with a slight abuse of notation we will still denote with $Z_\beta(\cdot)$. Then, to compare $\mathbf{G}_0$ and $\mathbf{G}_1$ we essentially show that

$$\frac{\partial}{\partial t}\mathbb{E}[\log Z_\beta(\mathbf{G}_t)] > 0,$$

so that $\mathbb{E}[\log Z_\beta(\mathbf{G}_0)]$ is upper-bounded by $\mathbb{E}[\log Z_\beta(\mathbf{G}_1)]$. Thus we can then directly compare $\mathbf{G}$ and $\mathbf{G}_1$. This kind of argument is generally known as the *interpolation method*.

The relation between partition functions of $\mathbb{G}$ and $\mathbf{G}$ was spelt out in Proposition 4.3.2, the relation between $\mathbf{G}$ and $\mathbf{G}_0$ is spelt out in Proposition 4.3.4, the relation between $\mathbf{G}_0$ and $\mathbf{G}_1$ in Proposition 4.3.3, and finally Proposition 4.3.5 gives a bound on the partition function of $\mathbf{G}_1$. Then Corollary 4.3.6 gives a formula for a bound on the partition function of the true d-regular graph $\mathbb{G}$, which only depends on the parameters β, γ and $\mathfrak{r}$ of the interpolation scheme defined shortly below. Corollary 4.3.6 and all Propositions in this section hold for any choice of parameters in the interpolation scheme. In Section 4.4 we prove Theorem 4.2.1 by making the suitable choices to these parameters.

First, observe that we can rewrite eqs. (4.7) and (4.8) as

$$\mu_{G,\beta}(\sigma) = \frac{1}{Z_\beta(G)}\prod_{uv\in E(G)} 1-(1-\exp(-\beta))\mathbf{1}\{\sigma_u=\sigma_v\}, \tag{4.13}$$

$$Z_\beta(G) = \sum_{\sigma\in\{\pm1\}^V}\prod_{uv\in E(G)} 1-(1-\exp(-\beta))\mathbf{1}\{\sigma_u=\sigma_v\}. \tag{4.14}$$

We denote with $\mathcal{P}(A)$ probability measures over a set A, and with $\mathcal{P}^2(A)$ the probability measures over the probability measures over A, and so on. In order to construct the interpolation scheme (the random graphs $(\mathbf{G}_t)_{t\in[0,1]}$), we first fix parameters $\beta,\varepsilon>0$, a probability measure $\mathfrak{r}\in\mathcal{P}^3(\{\pm1\})$, and a probability measure γ on $\mathbb{N}$. Let $(\mathbf{r}_i)_{i\geq1}$ be independent samples from $\mathfrak{r}$; thus, $\mathbf{r}_i\in\mathcal{P}^2(\{\pm1\})$. Further, for any $i,j\geq1$ sample $(\boldsymbol{\rho}_{i,h},\boldsymbol{\rho}_{i,j,h},\boldsymbol{\rho}'_{i,h},\boldsymbol{\rho}''_{i,h})_{i,j,h\geq1}$ independently such that $\boldsymbol{\rho}_{i,h},\boldsymbol{\rho}_{i,j,h},\boldsymbol{\rho}'_{i,h},\boldsymbol{\rho}''_{i,h}\in\mathcal{P}(\{\pm1\})$ have distribution $\mathbf{r}_i$ for all $h, j \geq 1$. Finally, let $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots)$, and let

$$\mathbf{M}_t \sim \mathrm{Po}\left((1-\varepsilon)(1-t)\frac{dn}{2}\right),\ \mathbf{M}'_t \sim \mathrm{Po}((1-\varepsilon)tdn),\ \mathbf{M}''_t \sim \mathrm{Po}\left((1-\varepsilon)(1-t)\frac{dn}{2}\right) \quad (4.15)$$

be independent Poisson random variables. Define the event

$$\mathcal{M} = \{2\mathbf{M}_t + \mathbf{M}' \leq dn,\ \mathbf{M}_t + \mathbf{M}'_t + \mathbf{M}''_t \leq dn, \mathbf{M}''_t \leq dn/2\} \quad (4.16)$$

and write $(\mathbf{m}_t, \mathbf{m}'_t, \mathbf{m}''_t)$ for $(\mathbf{M}_t, \mathbf{M}'_t, \mathbf{M}''_t)$ conditioned on $\mathcal{M}$. Note that $\mathbb{P}(\mathcal{M}) = 1 - \exp(-\Omega(n))$. The variables $\mathbf{m}_t$, $\mathbf{m}'_t$ and $\mathbf{m}''_t$ will denote the number of different types of constraint nodes in $\mathbf{G}_t$, defined shortly.

We also need the notion of a factor graph. A factor graph is represented by a bipartite graph with vertex classes *variable nodes V* and *constraint nodes C*. Each variable node can be assigned a spin, and each constraint node (and a special variable node s) will be assigned a weight that depends only on the spins of variable nodes adjacent to it. We let $\mathbf{G}_t$ for $t \in [0,1]$ be a factor graph variable nodes

$$s, v_1, \dots, v_n$$

and constraint nodes

$$e_1, \dots, e_{\mathbf{m}_t},\ a_1, \dots, a_{\mathbf{m}'_t},\ b_1, \dots, b_{\mathbf{m}''_t}.$$

Let $V_n = \{v_1, v_2, \dots, v_n\}$, so that the variable nodes of $\mathbf{G}_t$ are $\{s\} \cup V_n$. We construct $\mathbf{G}_t$ with the following configuration model. Let $\mathbf{L}_t$ be a random maximum matching of the complete bipartite graph with vertex classes

$$\left(\bigcup_{i=1}^{\mathbf{m}_t} \{e_i\} \times \{1,2\}\right) \cup \bigcup_{i=1}^{\mathbf{m}'_t} \{a_i\} \quad \text{and} \quad \left(\bigcup_{i=1}^{n} \{v_i\} \times [d]\right).$$

The matching $\mathbf{L}_t$ is left-saturated, as $2\mathbf{m}_t + \mathbf{m}'_t \leq dn$ (recall eq. (4.16)). We define $\mathbf{G}_t$ as follows:

1. Each constraint node a_i is adjacent to variable node s and variable node $u \in V_n$ for which there is an edge between a_i and $\{u\} \times [d]$ in $\mathbf{L}_t$.

2. Each constraint node e_i is adjacent to variable nodes $u, v \in V_n$ for which there is an edge between $(e_i, 1)$ and $\{u\} \times [d]$, and an edge between $(e_i, 2)$ and $\{v\} \times [d]$.

3. Each constraint node b_i is adjacent to variable node s only.

There are no other edges present in $\mathbf{G}_t$.

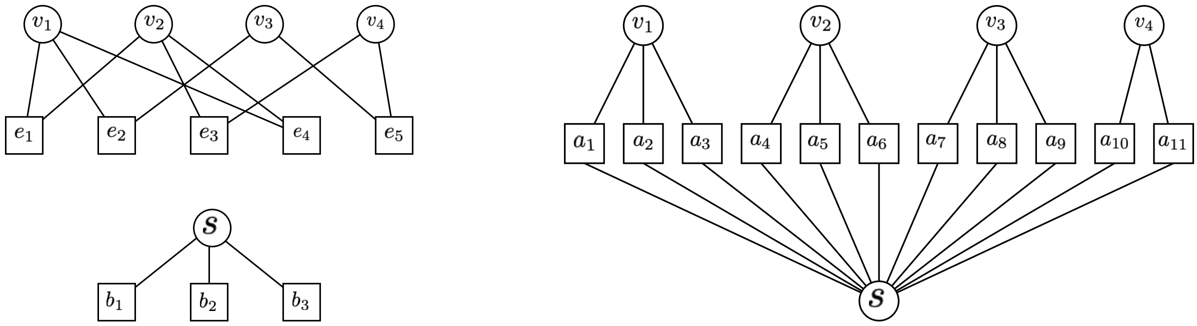


Fig. 4.1 The factor graph $\mathbf{G}_0$ (left) consisting of the original graph $\mathbf{G}$ and an auxiliary graph, and the factor graph $\mathbf{G}_1$ (right).

A *configuration* σ on a factor graph assigns each variable node $u \in V_n$ a spin in $\{\pm 1\}$ and the node s a value in $\mathbb{N}$. Thus $\sigma_s \in \mathbb{N}$ and $\sigma_u \in \{\pm 1\}$ for $u \in V_n$. Let $\Omega = \mathbb{N} \times \{\pm 1\}^{V_n}$ be the set of all configurations.

We define the following weights on the constraint nodes and on s depending on a configuration σ. Let

$$\begin{aligned}
\psi_s(\sigma_s) &= \gamma(\sigma_s) \\
\psi_{e_i}(\sigma_u, \sigma_v) &= 1 - (1 - e^{-\beta})\mathbf{1}\{\sigma_u = \sigma_v\} && (\partial e_i = \{u, v\}) \\
\psi_{a_i}(\sigma_s, \sigma_u) &= 1 - (1 - e^{-\beta})\boldsymbol{\rho}_{i,\sigma_s}(\sigma_u) && (\partial a_i = \{s, u\}) \\
\psi_{b_i}(\sigma_s) &= 1 - (1 - e^{-\beta}) \sum_{\tau \in \{\pm 1\}} \boldsymbol{\rho}'_{i,\sigma_s}(\tau)\boldsymbol{\rho}''_{i,\sigma_s}(\tau) && (\partial b_i = \{s\})
\end{aligned}$$

Thus, ψ_s weighs s according to the probability distribution γ; ψ_{e_i} model the edge weights assigned in the Ising model (compare to eq. (4.14)); ψ_{a_i} weighs the adjacent variable node according to $\boldsymbol{\rho}_{i,\sigma_s}$; ψ_{b_i} is determined by the probability that two spins sampled from $\boldsymbol{\rho}'_{i,\sigma_s}$ and $\boldsymbol{\rho}''_{i,\sigma_s}$ coincide, i.e. the probability that an edge between them is unsatisfied.

We define the partition function of $\mathbf{G}_t$ as the product of the weights defined above:

$$Z_\beta(\mathbf{G}_t) = \sum_{\sigma \in \Omega} \left(\psi_s(\sigma_s) \prod_{i=1}^{\mathbf{m}_t} \psi_{e_i}(\sigma_{\partial e_i}) \prod_{i=1}^{\mathbf{m}'_t} \psi_{a_i}(\sigma_{\partial a_i}) \prod_{i=1}^{\mathbf{m}''_t} \psi_{b_i}(\sigma_s) \right).$$

At 'time' $t = 0$, we have $\mathbf{m}'_0 = 0$ from eq. (4.15). Thus in $\mathbf{G}_0$ only constraint nodes e_i and b_i are present. See Fig. 4.1. Hence $\mathbf{G}_0$ is comprised of two parts: the component of node s which is a star with s at the centre connected to constraint nodes b_i, and the rest of the graph consists of V_n and constraint nodes e_i. Crucially, the subgraph induced by $v_1, \dots v_n$ and $e_1, \dots e_{\mathbf{m}'_t}$ is effectively identical to $\mathbf{G}$, as $\mathbf{m}$ (from eq. (4.11)) has the same distribution as $\mathbf{m}_0$. (As for $t = 0$, the event $\mathcal{M}$ reduces to event $\{\mathbf{M}_0 \le dn/2, \mathbf{M}''_0 \le dn/2\}$.) Further, as ψ_{e_i} mimic the edge penalties in $Z_\beta(\mathbf{G})$, we can relate the partition functions $Z_\beta(\mathbf{G})$ and $Z_\beta(\mathbf{G}_0)$, which we do Proposition 4.3.4.

At time $t = 1$, we have $\mathbf{m}_1 = \mathbf{m}''_1 = 0$. Thus in $\mathbf{G}_1$ only the constraint nodes a_i are present.

The following proposition spells out the relation between $\mathbf{G}_0$ and $\mathbf{G}_1$ and is the heart of the interpolation argument.

Proposition 4.3.3 ([ACG19, Proposition 2.8])**.** *We have*

$$\mathbb{E}[\log Z_\beta(\mathbf{G}_0)] \le \mathbb{E}[\log Z_\beta(\mathbf{G}_1)] + o(n). \tag{4.17}$$

As noted earlier, the factor graph $\mathbf{G}_0$ has two disjoint parts. Thus, the partition function $Z_\beta(\mathbf{G}_0)$ factorises as

$$Z_\beta(\mathbf{G}_0) = \mathcal{Y} \cdot \mathcal{Z}$$

where

$$\mathcal{Y} = \sum_{i=1}^{\infty} \gamma(\sigma_s) \prod_{i=1}^{\mathbf{m}''_0} \psi_{b_i}(\sigma_s), \quad \mathcal{Z} = \sum_{\sigma \in \{\pm 1\}^{V_n}} \prod_{i=1}^{\mathbf{m}_0} \psi_{e_i}(\sigma_{\partial e_i}).$$

Thus

$$\mathbb{E}[\log Z_\beta(\mathbf{G}_0)] = \mathbb{E}[\log \mathcal{Z}] + \mathbb{E}[\log \mathcal{Y}].$$

Further, by construction

$$\mathbb{E}[\log Z_\beta(\mathbf{G})] = \mathbb{E}[\log \mathcal{Z}],$$

and hence

$$\mathbb{E}[\log Z_\beta(\mathbf{G}_0)] = \mathbb{E}[\log Z_\beta(\mathbf{G})] + \mathbb{E}[\log \mathcal{Y}]. \tag{4.18}$$

The function (random variable) $\mathcal{Y}$ corresponds to the partition function of the s-component of $\mathbf{G}_0$. Define the random variable

$$Y' := \sum_{\sigma_s=1}^{\infty} \gamma(\sigma_s) \prod_{i=1}^{dn/2} 1 - (1 - e^{-\beta}) \sum_{\tau \in \{\pm 1\}} \rho'_{i,\sigma_s}(\tau) \rho''_{i,\sigma_s}(\tau). \tag{4.19}$$

Observe that Y' has the same distribution as $\mathcal{Y}$ if $\mathbf{m}''_0 = dn/2$. The following Proposition says that Y' is a good approximation of $\mathcal{Y}$. (Compare to eq. (4.18).)

Proposition 4.3.4 ([ACG19, Proposition 2.7]). *Let $\beta, \delta > 0$. Then for sufficiently small $\varepsilon > 0$,*

$$\mathbb{E}[\log Z_\beta(\mathbf{G}_0)] \geq \mathbb{E}[\log Z_\beta(\mathbf{G})] + \mathbb{E}[\log Y'] - \delta n.$$

We now turn our attention to $\mathbf{G}_1$. Define the random variable

$$Y := \sum_{\sigma_s=1}^{\infty} \gamma(\sigma_s) \prod_{i=1}^{n} \sum_{\tau \in \{\pm 1\}} \prod_{h=1}^{d} 1 - (1 - e^{-\beta}) \boldsymbol{\rho}_{i,\sigma_s,h}(\tau). \tag{4.20}$$

Observe that Y has the same distribution as $Z_\beta(\mathbf{G}_1)$ conditioned on $\mathbf{m}_1' = dn$. Again, we can use Y as an intermediate form to the partition function of $\mathbf{G}_1$.

Proposition 4.3.5 ([ACG19, Proposition 2.9]). *Let $\beta, \delta > 0$. Then for sufficiently small $\varepsilon > 0$,*

$$\mathbb{E}[\log Z_\beta(\mathbf{G}_1)] \leq \mathbb{E}[\log Y] + \delta n.$$

We are ready to give a bound on the partition function of $\mathbb{G}$ in terms of Y and Y', thus it only depends on the parameters β, γ and $\mathfrak{r}$ of the interpolation scheme.

Corollary 4.3.6. *For any $\beta > 0$, $\gamma \in \mathcal{P}(\mathbb{N})$ and $\mathfrak{r} \in \mathcal{P}^3(\{\pm 1\})$*

$$\mathbb{E}[\log Z_\beta(\mathbb{G})] \leq \mathbb{E}[\log Y] - \mathbb{E}[\log Y'] + o(n).$$

Proof. Immediate from Propositions 4.3.2 to 4.3.5. □

4.3.2 Poisson-Dirichlet weights

Still following [ACG19].

There is a choice of distribution for γ that greatly simplifies the formulas for $\mathbb{E}[\log Y]$ and $\mathbb{E}[\log Y']$. The *Dirichlet distribution with parameter $y > 0$* is defined as follows. Let $(\mathbf{x}_i)_{i\in\infty}$ be the sequence of points obtained from a Poisson Point Process $\mathbf{P}$ on $(0, \infty)$ with density $\mu = x^{-1-y}dx$, in decreasing order: $\mathbf{x}_1 > \mathbf{x}_2 > \dots$. As $y > 0$, we have that $\sum_{i=1}^{\infty} \mathbf{x}_i < \infty$ almost surely. We can then define the distribution $\boldsymbol{\gamma}$ on $\mathbb{N}$ by

$$\boldsymbol{\gamma}(s) = \frac{\mathbf{x}_s}{\sum_{i=1}^{\infty} \mathbf{x}_i}, \tag{4.21}$$

so that γ is a random probability distribution that depends on the Poisson Point Process $\mathbf{P}$.

Lemma 4.3.7 ([Tal03, Proposition 6.5.15]). *Let $(X_s)_{s\in\mathbb{N}}$ be positive i.i.d. random variables with $\mathbb{E}[X_s^2] < \infty$. Then, with γ as above,*

$$\mathbb{E}\big[\log \sum_{s=1}^{\infty} \gamma(s) X_s\big] = \frac{1}{y} \log \mathbb{E}[X_1^y].$$

Recall the definition of $\Phi_d(\beta)$ (our key quantity of interest) from eq. (4.12), and that $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots)$. So far we have omitted all proofs, but we give the proof of the following Corollary, to shed some light on the origins the formulas.

Corollary 4.3.8 ([SSZ16]). *For any $y, \beta > 0$ and $\mathfrak{r} \in \mathcal{P}^3(\{\pm 1\})$, we have $\Phi_d(\beta) \leq \phi_{\beta,y}(\mathfrak{r})$, where*

$$\phi_{\beta,y}(\mathfrak{r}) = \frac{1}{y}\mathbb{E}[\log \mathbb{E}[X^y \mid \mathbf{R}]] - \frac{d}{2y}\mathbb{E}[\log \mathbb{E}[X'^y \mid \mathbf{R}]] \tag{4.22}$$

with

$$\begin{aligned} X &= \sum_{\tau\in\{\pm1\}} \prod_{h=1}^{d} 1 - (1 - \mathrm{e}^{-\beta}) \rho_{1,1,h}(\tau), \\ X' &= 1 - (1 - \mathrm{e}^{-\beta}) \sum_{\tau\in\{\pm1\}} \rho'_{1,1}(\tau) \rho''_{1,1}(\tau). \end{aligned}$$

Proof. Define the random variables,

$$\begin{aligned} X_k &= \prod_{i=1}^{n} \sum_{\tau\in\{\pm1\}} \prod_{h=1}^{d} 1 - (1 - \mathrm{e}^{-\beta}) \rho_{i,k,h}(\tau), \\ X'_k &= \prod_{i=1}^{dn/2} 1 - (1 - \mathrm{e}^{-\beta}) \sum_{\tau\in\{\pm1\}} \rho'_{i,k}(\tau) \rho''_{i,k}(\tau), \end{aligned}$$

Thus, we have that

$$Y = \sum_{k=1}^{\infty} \gamma(k) X_k \quad \text{and} \quad Y' = \sum_{k=1}^{\infty} \gamma(k) X'_k. \tag{4.23}$$

Applying Corollary 4.3.6, Lemma 4.3.7, and eq. (4.23) with $\gamma = \boldsymbol{\gamma}$ (which depends on the Poisson Point Process $\mathbf{P}$, eq. (4.21)), we obtain

$$\begin{aligned}\mathbb{E}[\log Z_\beta(\mathbb{G})] &\leq \mathbb{E}[\log Y] - \mathbb{E}[\log Y'] + o(n)\\ &= \mathbb{E}\left[\log\left(\sum_{k=1}^{\infty} \boldsymbol{\gamma}(k) X_k\right)\right] - \mathbb{E}\left[\log\left(\sum_{k=1}^{\infty} \boldsymbol{\gamma}(k) X'_k\right)\right] + o(n)\\ &= \frac{1}{y}\,\mathbb{E}[\log \mathbb{E}[X_1^y \mid \mathbf{R}]] - \frac{1}{y}\,\mathbb{E}[\log \mathbb{E}[X_1'^y \mid \mathbf{R}]] + o(n).\end{aligned}$$

As $(\boldsymbol{\rho}_{i,k,h}, \boldsymbol{\rho}'_{i,k}, \boldsymbol{\rho}''_{i,k})_{k,h\geq 1}$ are independent given $\mathbf{R}$, both $\mathbb{E}[X_1^y]$ and $\mathbb{E}[X_1'^y]$ factorise. Hence,

$$\begin{aligned}\mathbb{E}[\log \mathbb{E}[X_1^y \mid \mathbf{R}]] &= n\,\mathbb{E}[\log \mathbb{E}[X^y \mid \mathbf{R}]],\\ \mathbb{E}[\log \mathbb{E}[X_1'^y \mid \mathbf{R}]] &= \frac{dn}{2}\,\mathbb{E}[\log \mathbb{E}[X'^y \mid \mathbf{R}]],\end{aligned}$$

completing the proof.

□

4.4 Proof of main result

We now return to proving Theorem 4.2.1, which is our contribution.

It remains to choose the correct parameters β, y and $\mathfrak{r} \in \mathcal{P}^3(\{\pm 1\})$ in Corollary 4.3.6. We take the limit $\beta \to \infty$, which in physical terms corresponds to the zero temperature limit. We will take the limit such that $\beta y = \Theta(1)$, see eq. (4.26) below. Similar limits were taken in [DSS15] to derive the upper bound on the k-SAT threshold from the formula for the k-SAT partition function from [PT04], and in [ACG19] to derive a lower bound on the chromatic number of random regular graphs.

For $i \in \{\pm 1\}$, let $\delta_i \in \mathcal{P}(\{\pm 1\})$ be the atom on spin +1 and -1, respectively. Then for $p \in \{0, 1/2, 1\}$ we define

$$r_p = p\delta_{\nu_{-1}} + (1-p)\delta_{\nu_{+1}} \in \mathcal{P}^2(\{\pm 1\}). \tag{4.24}$$

Thus r_p is a distribution over distributions on $\{\pm 1\}$, which with probability (w.p.) p makes the spin +1 (w.p. 1), and w.p. $1-p$ makes the spin -1 (w.p. 1). Further, for $\alpha \in (0, 1/2)$ let

$$\mathfrak{r}_\alpha = \alpha r_0 + (1-2\alpha) r_{1/2} + \alpha r_1 \in \mathcal{P}^3(\{\pm 1\}) \tag{4.25}$$

Intuitively, we can think of $\mathfrak{r}_\alpha$ as setting a vertex to a guaranteed +1 and -1 with probability α each, and randomly selecting between +1 and −1 otherwise.

In the following we use the substitution

$$y = -\log(z)/\beta \tag{4.26}$$

for a fixed $z > 0$.

Lemma 4.4.1. *For $\alpha \in (0, 1/2)$, $z \in (0, 1)$, we have*

$$\lim_{\beta\to\infty} \mathbb{E}[\log \mathbb{E}[X'^y \mid \mathbf{R}]] = \log\left(1 - 2\alpha^2 + 2\alpha^2 z\right).$$

Proof. In the following, we abbreviate $\left(1-(1-e^{-\beta})(\rho_1\rho_2+(1-\rho_1)(1-\rho_2))\right)^y$ as $\mathcal{E}$. Then,

$$\begin{aligned}
\lim_{\beta\to\infty} \mathbb{E}[\log \mathbb{E}[X'^y \mid \mathbf{R}]] &= \lim_{\beta\to\infty} \mathbb{E}\left[\log \mathbb{E}\left[\left(1-(1-\mathrm{e}^{-\beta})\sum_{\tau\in\{\pm1\}} \boldsymbol{\rho}'_{1,1}(\tau)\boldsymbol{\rho}''_{1,1}(\tau)\right)^y \,\middle|\, \mathbf{R}\right]\right] \\
&= \lim_{\beta\to\infty} \log\left(\sum_{\rho_1,\rho_2\in\{0,1/2,1\}} \mathfrak{r}_\alpha(r_{\rho_1})\mathfrak{r}_\alpha(r_{\rho_2}) \cdot \mathcal{E}^y\right) \\
&= \log\left(1-2\alpha^2+2\alpha^2 z\right).
\end{aligned}$$

The first equality is by definition of X'. For the second equality, we look to eq. (4.25). In the inner expectation, we are sampling $\boldsymbol{\rho}'_{1,1}$ and $\boldsymbol{\rho}''_{1,1}$ according to $\mathbf{r}_1$, and therefore we can simply write this as the sum over the 9 possible outcomes. The term $\mathfrak{r}_\alpha(r_{\rho_1})\mathfrak{r}_\alpha(r_{\rho_2})$ corresponds to the probability of the given outcome, and $\mathcal{E}^y$ is the value given by the inner expectation in this case. The outer expectation disappears, as the formula no longer depends on $\mathbf{r}_1$ (or $\mathbf{R}$).

To understand the third (asymptotic) equality, observe that there are four possibilities for the relation between ρ_1 and ρ_2.

- If $\rho_1, \rho_2 \in \{0, 1\}$ and $\rho_1 \neq \rho_2$, which occurs with probability $2\alpha^2$, $\mathcal{E}$ simplifies to $1^y = 1$.
- If $\rho_1, \rho_2 \in \{0, 1\}$ and $\rho_1 = \rho_2$, which occurs with probability $2\alpha^2$, $\mathcal{E} = \exp(-\beta y)$.
- If $\rho_1 \in \{0, 1\}$ and $\rho_2 = 1/2$ or vice versa, which occurs with probability $4\alpha(1 - 2\alpha)$, $\mathcal{E} = ((1 + e^{-\beta})/2)^y \to 1$, as $y \to 0, \beta \to \infty$.
- If $\rho_1, \rho_2 = 1/2$, which occurs with probability $(1 - 2\alpha)^2$, $\mathcal{E} = ((1 + e^{-\beta})/2)^y \to 1$ as before.

Therefore, with probability $1-2\alpha^2$, $\mathcal{E}$ tends to 1, while it takes the value $z = \exp(-\beta y)$ in the remaining cases. Intuitively, a penalty term z is therefore only added, if both vertices connected by an edge surely take the same spin. $\square$

Deriving an explicit bound for $\lim_{\beta\to\infty} \mathbb{E}[\log \mathbb{E}[X^y \mid \mathbf{R}]]$ is more involved.

Lemma 4.4.2. *For $\alpha \in (0, 1/2)$, $z \in (0, 1)$, we have*

$$\lim_{\beta\to\infty} \mathbb{E}\left[\log \mathbb{E}\left[X^y \mid \mathbf{R}\right]\right] = \log\left(\zeta \mathcal{A}^d_{d,\alpha} \xi|_{t=z^{1/2}}\right).$$

Proof. Consider the inner term

$$X^y = \left(\sum_{\tau\in\{\pm 1\}} \prod_{h=1}^{d} 1 - (1-\mathrm{e}^{-\beta})\boldsymbol{\rho}_{1,1,h}(\tau)\right)^y,$$

and define the following random variables:

$$\mathbf{A} = \sum_{h=1}^{d} \mathbf{1}\{\boldsymbol{\rho}_{1,1,h}(1) = 0\}, \quad \mathbf{B} = \sum_{h=1}^{d} \mathbf{1}\{\boldsymbol{\rho}_{1,1,h}(1) = 1/2\}, \quad \mathbf{C} = \sum_{h=1}^{d} \mathbf{1}\{\boldsymbol{\rho}_{1,1,h}(1) = 1\}.$$

Note that $\mathbf{A} \sim \mathrm{Bin}(d, \alpha)$, $\mathbf{C} \sim \mathrm{Bin}(d, \alpha)$, $\mathbf{B} \sim \mathrm{Bin}(d, 1-2\alpha)$ conditioned on $\mathbf{A}+\mathbf{B}+\mathbf{C} = d$. Further,

$$\left(\prod_{h=1}^{d} 1 - (1-e^{-\beta})\boldsymbol{\rho}_{1,1,h}(1)\right)^y = \exp(-\beta y \mathbf{C})\left(\frac{1+e^{-\beta}}{2}\right)^{y\mathbf{B}} \sim \exp\left(-\beta y \mathbf{C}\right) \text{ as } \beta \to \infty,$$
$$\left(\prod_{h=1}^{d} 1 - (1-e^{-\beta})\boldsymbol{\rho}_{1,1,h}(-1)\right)^y = \exp(-\beta y \mathbf{A})\left(\frac{1+e^{-\beta}}{2}\right)^{y\mathbf{B}} \sim \exp(-\beta y \mathbf{A}) \text{ as } \beta \to \infty.$$

Thus,

$$X^y \sim \exp(-\beta y(\mathbf{A} \wedge \mathbf{C})) \quad \text{as } \beta \to \infty.$$

Finally, using $z = \exp(-\beta y)$, we obtain

$$\lim_{\beta\to\infty} \mathbb{E}[\log \mathbb{E}[X^y \mid \mathbf{R}]] = \mathbb{E}\left[\log \mathbb{E}\left[z^{\mathbf{A}\wedge\mathbf{C}} \mid \mathbf{R}\right]\right].$$

To calculate this consider a d-step symmetric random walk on $\mathbb{N}$ with a reflective barrier at 0. At $k \in \mathbb{N}$, for $k \geq 1$ the available moves are $+1$, -1 or 0, with probabilities α, α

and $1-2\alpha$ respectively; and for $k=0$ the available moves are $+1$ or 0 with probabilities 2α and $1-2\alpha$ respectively. Let $\mathbf{A}$ and $\mathbf{C}$ count the $+1$ and -1 moves, respectively, and $\mathbf{B}$ the 0 moves. Thus, $\mathbf{A}+\mathbf{C}$ is the total number of non-stationary moves and $|\mathbf{A}-\mathbf{C}|$ the final position of the walk. Consider the matrix

$$\mathcal{A}_{d,\alpha} = (1-2\alpha)\mathrm{id} + 2\alpha t \mathcal{M}_d,$$

where t is a dummy variable that we introduce to count upwards and downwards movements and $\mathcal{M}_d$ is defined as in eq. (4.1). Thus for every $i \in [d]$,

$$\left(\mathcal{A}_{d,\alpha}^d\right)_{1\,i} = \sum_{k=i-1}^{d} t^k \mathbb{P}\left[\mathbf{A}+\mathbf{C}=k \text{ and } |\mathbf{A}-\mathbf{C}| = i-1\right].$$

Further, define vectors ζ and ξ as in eq. (4.2) and eq. (4.3). Consider the matrix product $\zeta\mathcal{A}_{d,\alpha}^d\xi$ and observe that the exponent of t equals $\mathbf{A}+\mathbf{C}-|\mathbf{A}-\mathbf{C}| = 2(\mathbf{A}\wedge\mathbf{C})$. Therefore, substituting $t=\sqrt{z}$ yields

$$\mathbb{E}\left[\log\mathbb{E}\left[z^{\mathbf{A}\wedge\mathbf{C}} \mid \mathbf{R}\right]\right] = \log\left(\zeta\mathcal{A}_{d,\alpha}^d\xi\big|_{t=z^{1/2}}\right).$$

□

Recall the definition of $F_d(\alpha,z)$ from eq. (4.4) and that we are using the substitution eq. (4.26). Corollary 4.3.8 and Lemmas 4.4.1 and 4.4.2 yield

$$\limsup_{\beta\to\infty}\frac{\Phi_d(\beta)}{\beta} \le F_d(\alpha,z) \tag{4.27}$$

for all $\alpha\in(0,1/2), z\in(0,1)$.

We are now ready to prove our main theorem.

Proof of Theorem 4.2.1. By applying eq. (4.9) to $\mathbb{G}$, for any $\beta>0$ we have

$$\frac{2}{dn}\mathbb{E}[\mathrm{MaxCut}(\mathbb{G})] \le 1 + \frac{2}{\beta dn}\mathbb{E}[\log Z_\beta(\mathbb{G})].$$

Hence, using eq. (4.27),

$$\limsup_{n\to\infty}\frac{2}{dn}\mathbb{E}[\mathrm{MaxCut}(\mathbb{G})] \le 1 + \frac{2}{d}\limsup_{\beta\to\infty}\frac{\Phi_d(\beta)}{\beta} \le 1 + \frac{2}{d}\inf_{\substack{\alpha\in(0,1/2)\\ z\in(0,1)}} F_d(\alpha,z). \tag{4.28}$$

The assertion eq. (4.5) follows from eq. (4.28) and Proposition 4.3.1. □

4.5 Numerical data

Values in Table 4.1 (for Corollary 4.2.2) are obtained from the following α and z values. The numerical optimization has been done using the *Optim.jl* numerical optimization package [KN18] written in the *Julia* programming language [BEKS17].

d	$F(\alpha, z)$	α	z
3	-0.11385	0.39104	0.43395
4	-0.26353	0.47244	0.45909
5	-0.41240	0.42992	0.52974
6	-0.58541	0.47264	0.53430
7	-0.75217	0.44600	0.58249
8	-0.93659	0.47382	0.58312
9	-1.11453	0.45511	0.61837
10	-1.30615	0.47513	0.61815

Table 4.2 Optimal numerical values for $F_d(\alpha, z)$ for $d = 3 \dots 10$

References

[ACG19] P. Ayre, A. Coja-Oghlan, and C. Greenhill. “Lower Bounds on the Chromatic Number of Random Graphs”. June 26, 2019. arXiv: 1812.09691.

[AJ05] J. A. Adell and P. Jodrá. “The Median of the Poisson Distribution”. In: *Metrika* 61.3 (June 2005), pp. 337–346. ISSN: 0026-1335, 1435-926X. DOI: 10.1007/s001840400350.

[AK95] C. J. Alpert and A. B. Kahng. “Recent Directions in Netlist Partitioning: A Survey”. In: *Integration* 19.1-2 (Aug. 1995), pp. 1–81. ISSN: 01679260. DOI: 10.1016/0167-9260(95)00008-4.

[Ald01] D. J. Aldous. “The $\zeta(2)$ Limit in the Random Assignment Problem”. In: *Random Structures and Algorithms* 18.4 (July 2001), pp. 381–418. ISSN: 1042-9832, 1098-2418. DOI: 10.1002/rsa.1015.

[Alo97] N. Alon. “On the Edge-Expansion of Graphs”. In: *Combinatorics, Probability and Computing* 6.2 (June 1997), pp. 145–152. ISSN: 09635483. DOI: 10.1017/S096354839700299X.

[ARV04] S. Arora, S. Rao, and U. Vazirani. “Expander Flows, Geometric Embeddings and Graph Partitioning”. In: *Journal of the ACM* 56 (Apr. 30, 2004). DOI: 10.1145/1502793.1502794.

[BCLS87] T. N. Bui, S. Chaudhuri, F. T. Leighton, and M. Sipser. “Graph Bisection Algorithms with Good Average Case Behavior”. In: *Combinatorica* 7.2 (June 1, 1987), pp. 171–191. ISSN: 1439-6912. DOI: 10.1007/BF02579448.

[BEKS17] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. “Julia: A Fresh Approach to Numerical Computing”. In: *SIAM Review* 59.1 (Jan. 2017), pp. 65–98. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/141000671.

[BFM98] A. Beveridge, A. Frieze, and C. McDiarmid. “Random Minimum Length Spanning Trees in Regular Graphs”. In: *Combinatorica* 18.3 (Mar. 1, 1998), pp. 311–333. ISSN: 1439-6912. DOI: 10.1007/PL00009825.

[BH12] S. Bhamidi and R. van der Hofstad. “Weak Disorder Asymptotics in the Stochastic Mean-Field Model of Distance”. In: *Annals of Applied Probability* 22.1 (Feb. 2012), pp. 29–69. ISSN: 1050-5164, 2168-8737. DOI: 10.1214/10-AAP753.

[BHH11] S. Bhamidi, R. Hofstad, and G. Hooghiemstra. “First Passage Percolation on the Erdős-Rényi Random Graph”. In: *Combinatorics, Probability & Computing* 20 (Sept. 1, 2011), pp. 683–707. DOI: 10.1017/S096354831100023X.

[Bol88] B. Bollobás. "The Isoperimetric Number of Random Regular Graphs". In: *European Journal of Combinatorics* 9.3 (May 1988), pp. 241–244. ISSN: 01956698. DOI: 10.1016/S0195-6698(88)80014-3.

[Bol98] B. Bollobás. *Modern Graph Theory*. Vol. 184. Graduate Texts in Mathematics. Springer-Verlag, New York, 1998. DOI: 10.1007/978-1-4612-0619-4.

[BT97] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific Series in Optimization and Neural Computation. Athena Scientific, 1997. 587 pp.

[Cho94] K. P. Choi. "On the Medians of Gamma Distributions and an Equation of Ramanujan". In: *Proceedings of the American Mathematical Society* 121.1 (May 1994), p. 245. ISSN: 00029939. DOI: 10.2307/2160389.

[DHL17] I. Dunning, J. Huchette, and M. Lubin. "JuMP: A Modeling Language for Mathematical Optimization". In: *SIAM Review* 59.2 (Jan. 2017), pp. 295–320. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/15M1020575.

[Dij59] E. W. Dijkstra. "A Note on Two Problems in Connexion with Graphs". In: *Numerische Mathematik* 1.1 (Dec. 1959), pp. 269–271. ISSN: 0029-599X, 0945-3245. DOI: 10.1007/BF01386390.

[DMS17] A. Dembo, A. Montanari, and S. Sen. "Extremal Cuts of Sparse Random Graphs". In: *The Annals of Probability* 45.2 (Mar. 2017), pp. 1190–1217. ISSN: 0091-1798. DOI: 10.1214/15-AOP1084.

[DSS15] J. Ding, A. Sly, and N. Sun. "Proof of the Satisfiability Conjecture for Large k". In: *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing - STOC '15*. ACM Press, 2015, pp. 59–68. DOI: 10.1145/2746539.2746619.

[DSW07] J. Díaz, M. Serna, and N. Wormald. "Bounds on the Bisection Width for Random d-Regular Graphs". In: *Theoretical Computer Science* 382.2 (Aug. 2007), pp. 120–130. ISSN: 03043975. DOI: 10.1016/j.tcs.2007.03.003.

[EGvdHN13] M. Eckhoff, J. Goodman, R. van der Hofstad, and F. R. Nardi. "Short Paths for First Passage Percolation on the Complete Graph". In: *Journal of Statistical Physics* 151.6 (June 1, 2013), pp. 1056–1088. ISSN: 1572-9613. DOI: 10.1007/s10955-013-0743-7.

[FJ18] A. Frieze and T. Johansson. "On Edge-Disjoint Spanning Trees in a Randomly Weighted Complete Graph". In: *Combinatorics, Probability and Computing* 27.2 (Mar. 2018), pp. 228–244. ISSN: 0963-5483, 1469-2163. DOI: 10.1017/S0963548317000426.

[FPS18] A. Frieze, W. Pegden, and G. B. Sorkin. "The Distribution of Minimum-Weight Cliques and Other Subgraphs in Graphs with Random Edge Weights". In: *SIAM Journal on Discrete Mathematics* 32.3 (Jan. 2018), pp. 2115–2133. ISSN: 0895-4801, 1095-7146. DOI: 10.1137/17M1138303.

[Fri04] A. Frieze. "On Random Symmetric Travelling Salesman Problems". In: *Mathematics of Operations Research* 29.4 (Nov. 2004), pp. 878–890. ISSN: 0364-765X, 1526-5471. DOI: 10.1287/moor.1040.0105.

[Fri85] A. Frieze. "On the Value of a Random Minimum Spanning Tree Problem". In: *Discrete Applied Mathematics* 10.1 (Jan. 1985), pp. 47–56. ISSN: 0166218X. DOI: 10.1016/0166-218X(85)90058-7.

[GJ09] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. A Series of Books in the Mathematical Sciences. Freeman, 2009. 338 pp.

[GJS76] M. Garey, D. Johnson, and L. Stockmeyer. "Some Simplified NP-Complete Graph Problems". In: *Theoretical Computer Science* 1.3 (Feb. 1976), pp. 237–267. ISSN: 03043975. DOI: 10.1016/0304-3975(76)90059-1.

[GW95] M. X. Goemans and D. P. Williamson. "Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming". In: *Journal of the ACM* 42 (1995), pp. 1115–1145.

[Ham95] K. Hamza. "The Smallest Uniform Upper Bound on the Distance between the Mean and the Median of the Binomial and Poisson Distributions". In: *Statistics & Probability Letters* 23.1 (Apr. 1995), pp. 21–25. ISSN: 01677152. DOI: 10.1016/0167-7152(94)00090-U.

[HL95] B. Hendrickson and R. Leland. "A Multilevel Algorithm for Partitioning Graphs". In: *Proceedings of the 1995 ACM/IEEE Conference on Supercomputing (CDROM) - Supercomputing '95*. ACM Press, 1995, 28–es. DOI: 10.1145/224170.224228.

[Hla06] J. Hladký. "Structural Properties of Graphs — Probabilistic and Deterministic Point of View". Bachelor's Thesis. 2006.

[HW16] C. Hoppen and N. Wormald. "Properties of Regular Graphs with Large Girth via Local Algorithms". In: *Journal of Combinatorial Theory, Series B* 121 (Nov. 2016), pp. 367–397. ISSN: 00958956. DOI: 10.1016/j.jctb.2016.07.009.

[Jan02] S. Janson. "On Concentration of Probability". In: *Combinatorics, Probability and Computing* (2002).

[Jan18] S. Janson. "Tail Bounds for Sums of Geometric and Exponential Variables". In: *Statistics & Probability Letters* 135 (Apr. 2018), pp. 1–6. ISSN: 01677152. DOI: 10.1016/j.spl.2017.11.017.

[Jan99] S. Janson. "One, Two and Three Times Log n/n for Paths in a Complete Graph with Random Weights". In: *Combinatorics, Probability and Computing* 8.4 (July 1999), pp. 347–361. ISSN: 09635483. DOI: 10.1017/S0963548399003892.

[JS19] S. Janson and G. B. Sorkin. "Successive Minimum Spanning Trees". June 4, 2019. arXiv: 1906.01533.

[KB80] R. Kaas and J. Buhrman. "Mean, Median and Mode in Binomial Distributions". In: *Statistica Neerlandica* 34.1 (Mar. 1980), pp. 13–18. ISSN: 0039-0402, 1467-9574. DOI: 10.1111/j.1467-9574.1980.tb00681.x.

[KM92] A. Kostochka and L. Mel'nikov. "On Bounds of the Bisection Width of Cubic Graphs". In: *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity*. Ed. by J. Neŝetril and M. Fiedler. Vol. 51. Annals of Discrete Mathematics. Elsevier, 1992, pp. 151–154. DOI: 10.1016/S0167-5060(08)70620-4.

[KN18] P. K Mogensen and A. N Riseth. "Optim: A Mathematical Optimization Package for Julia". In: *Journal of Open Source Software* 3.24 (Apr. 4, 2018), p. 615. ISSN: 2475-9066. DOI: 10.21105/joss.00615.

[LLR95] N. Linial, E. London, and Y. Rabinovich. "The Geometry of Graphs and Some of Its Algorithmic Applications". In: *Combinatorica* 15.2 (June 1, 1995), pp. 215–245. ISSN: 1439-6912. DOI: 10.1007/BF01200757.

[LO18] M. Lelarge and M. Oulamara. "Replica Bounds by Combinatorial Interpolation for Diluted Spin Systems". In: *Journal of Statistical Physics* 173.3 (Nov. 1, 2018), pp. 917–940. ISSN: 1572-9613. DOI: 10.1007/s10955-018-1964-6.

[Luc92] É. Lucas. *Récréations Mathématiques*. A facsimile of the 1896 second printing is accessible at https://archive.org/details/recretionmatedou02lucarich. Gauthier-Villars, 1892.

[Mar16] A. Martinsson. "Unoriented First-Passage Percolation on the n-Cube". In: *Annals of Applied Probability* 26.5 (Oct. 2016), pp. 2597–2625. ISSN: 1050-5164, 2168-8737. DOI: 10.1214/15-AAP1155.

[McK82] B. McKay. "Maximum Bipartite Subgraphs of Regular Graphs with Large Grith". In: *Proceedings of the 13th Southeastern Conference on Combinatorics, Graph Theory and Computing*. Utilitas Mathematica Pub., 1982.

[MP06] B. Monien and R. Preis. "Upper Bounds on the Bisection Width of 3- and 4-Regular Graphs". In: *Journal of Discrete Algorithms* 4.3 (Sept. 2006), pp. 475–498. ISSN: 15708667. DOI: 10.1016/j.jda.2005.12.009.

[MPD00] B. Monien, R. Preis, and R. Diekmann. "Quality Matching and Local Improvement for Multilevel Graph-Partitioning". In: *Parallel Computing* 26.12 (Nov. 2000), pp. 1609–1634. ISSN: 01678191. DOI: 10.1016/S0167-8191(00)00049-1.

[Pan13] D. Panchenko. *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer New York, 2013. DOI: 10.1007/978-1-4614-6289-7.

[PT04] D. Panchenko and M. Talagrand. "Bounds for Diluted Mean-Fields Spin Glass Models". May 18, 2004. arXiv: math/0405357.

[PT93] S. Poljak and Z. Tuza. "Maximum Cuts and Largest Bipartite Subgraphs". In: *Combinatorial Optimization*. 1993. DOI: 10.1090/dimacs/020/04.

[Räc08] H. Räcke. "Optimal Hierarchical Decompositions for Congestion Minimization in Networks". In: Proc. 40th STOC. May 17, 2008, pp. 255–264. DOI: 10.1145/1374376.1374415.

[SFH06] C. Su, Q. Feng, and Z. Hu. "Uniform Recursive Trees: Branching Structure and Simple Random Downward Walk". In: *Journal of Mathematical Analysis and Applications* 315.1 (Mar. 2006), pp. 225–243. ISSN: 0022247X. DOI: 10.1016/j.jmaa.2005.05.004.

[Shm97] D. B. Shmoys. "Cut Problems and Their Application to Divide-and-Conquer". In: *Approximation Algorithms for NP-Hard Problems*. Ed. by D. S. Hochbaum. PWS Pub. Co, 1997, pp. 192–235.

[Sor20] G. B. Sorkin. "Extremal Cuts in Random Cubic Graphs". Manuscript (28 pages). 2020.

[SSZ16] A. Sly, N. Sun, and Y. Zhang. "The Number of Solutions for Random Regular NAE-SAT". In: *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, Oct. 2016, pp. 724–731. DOI: 10.1109/FOCS.2016.82.

[Tal03] M. Talagrand. *Spin Glasses: A Challenge for Mathematicians: Cavity and Mean Field Models*. Ergebnisse Der Mathematik Und Ihrer Grenzgebiete 3. Folge, v. 46. Springer, 2003. 586 pp.

[Tal06] M. Talagrand. "The Parisi Formula". In: *Annals of Mathematics* 163.1 (2006), pp. 221–263. ISSN: 0003-486X. JSTOR: 20159953.

[Wäs08] J. Wästlund. "Random Matching Problems on the Complete Graph". In: *Electronic Communications in Probability* 13.0 (2008), pp. 258–265. ISSN: 1083-589X. DOI: 10.1214/ECP.v13-1372.

[Wäs09] J. Wästlund. "An Easy Proof of the $\zeta(2)$ Limit in the Random Assignment Problem". In: *Electronic Communications in Probability* 14.0 (2009), pp. 261–269. ISSN: 1083-589X. DOI: 10.1214/ECP.v14-1475.

[Wäs10] J. Wästlund. "The Mean Field Traveling Salesman and Related Problems". In: *Acta Mathematica* 204.1 (Mar. 1, 2010), pp. 91–150. ISSN: 1871-2509. DOI: 10.1007/s11511-010-0046-7.

[ZB10] L. Zdeborová and S. Boettcher. "A Conjecture on the Maximum Cut and Bisection Width in Random Regular Graphs". In: *Journal of Statistical Mechanics: Theory and Experiment* 2010.02 (Feb. 25, 2010), P02020. ISSN: 1742-5468. DOI: 10.1088/1742-5468/2010/02/P02020.