Expected cost in Combinatorial Optimization under color constraints

Patrick Bennett* Alan Frieze† Wesley Pegden‡

August 3, 2026

Abstract

We present an average case model of classical problems in combinatorial optimization where there are color constraints. In all cases we seek some (spanning) sub-structure of a complete graph of minimum cost. The edges are randomly colored either red or blue. We bias against the red edges by placing a bound on the number of them that are allowed in our structure. This bound will be lower w.h.p. than what would occur without discrimination. We examine the effect of this bias on the minimum cost of a desired structure. We consider minimum cost spanning trees, shortest paths, minimum cost perfect matchings and the asymmetric traveling salesperson problem.

1 Introduction

We present an average case model of classical problems in combinatorial optimization where there are color constraints. In all cases we seek some (spanning) sub-structure of a complete graph of minimum cost. The edges are colored either red or blue. We bias against the red edges by placing a bound on the number of them that are allowed in our structure. This bound will be lower than what would occur with high probability without discrimination. We examine the effect of this bias on the minimum cost of a desired structure.

The analysis of randomly colored random graphs is a thriving area, see Chapter 15 of [7] for an overview of some of the best studied questions. In this paper we deal with random colorings within the context of optimization.

In this paper we use standard asymptotic notation, and asymptotics are as $n \to \infty$. We write $a_n \sim b_n$ if $a_n = (1 + o(1))b_n$. We say a sequence of events $\mathcal{E}_n$ occurs *with high probability (w.h.p.)* if $\mathbb{P}[\mathcal{E}_n] \to 1$.

*Department of Mathematics, Western Michigan University, Kalamazoo MI 49008-5248, Research supported in part by Simons Foundation Grant #848648. Email: patrick.bennett@wmich.edu.

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, Research supported in part by NSF grant DMS2341774. Email: frieze@cmu.edu.

‡Research supported in part by NSF grant DMS-1363136. Email: wes@math.cmu.edu.

1.1 Spanning Trees

First consider the complete graph K_n where the edges are colored either red or blue independently with probability 1/2. Each edge $e \in E(K_n)$ is given an independent uniform $[0,1]$ cost X_e. We let E_R denote the set of red edges. We seek a minimum cost spanning tree that contains at most αn red edges, where $\alpha < 1/2$. (We will assume that αn is an integer.) The cost $C(T)$ of a tree T is given by $\sum_{e\in T} X_e$. Note that when $\alpha > 1/2$ this restriction will be redundant w.h.p. Indeed, if we ignore the color constraint and just find the minimum cost spanning tree, w.h.p. it has at most αn red edges. Frieze [6] proved that with no color constraint, the minimum cost spanning tree has cost $\zeta(3) = \sum_{k=1}^{\infty} k^{-3}$ w.h.p. In this context we prove

Theorem 1. *Let $Z_\alpha = Z_\alpha(MST)$ denote the minimum cost of a spanning tree T with $|E_R \cap T| \le \alpha n$. If $\frac{1}{4} < \alpha < \frac{1}{2}$, we have*

$$\mathbb{E}Z_\alpha \sim \zeta(3) + (1-2\alpha)^2.$$

More generally, let

$$f(x) = \frac{1}{x}\sum_{k=1}^{\infty} \frac{k^{k-2}}{k!}(xe^{-x})^k = \begin{cases} 1 - \frac{x}{2} & x \le 1. \\ \frac{y}{x}\left(1 - \frac{y}{2}\right) & x > 1. \end{cases} \tag{1}$$

Here, $y < 1$ satisfies $ye^{-y} = xe^{-x}$.

If ψ is the solution to

$$f(\psi) = 2\alpha.$$

then for any $0 < \alpha < \frac{1}{2}$, we have

$$\mathbb{E}Z_\alpha \sim \zeta(3) + \int_{x=0}^{\psi} f(x)\,dx - 2\alpha\psi. \tag{2}$$

Furthermore, $Z_\alpha \sim \mathbb{E}Z_\alpha$ w.h.p.

The function $f(x)$ is asymptotically equal to n^{-1} times the expected number of components in the random graph $G = G_{n,x/n}$. If $x < 1$ then w.h.p. G is a forest plus a few unicyclic components and so there are close to $n - |E(G)|$ components, where $|E(G)| \sim nx/2$ w.h.p. We do not prove what happens when $\alpha = 1/2$, but letting $\alpha \nearrow 1/2$ into the above formula gives $\psi \to 0$ and $Z_\alpha \sim \zeta(3)$, which corresponds to the result of [6].

1.2 Paths

For this problem each edge e of the complete graph (or digraph) is given a cost X_e independently distributed as exponential mean 1, $EXP(1)$, i.e. $\mathbb{P}(X_e \ge x) = e^{-x}$ for $x \ge 0$. Edges are then independently colored red or blue with probability 1/2. Now for $0 < \alpha < 1/2$ we compute minimum cost paths under the restriction that

$$\text{no path can contain more than } L = \alpha \log n \text{ red edges.} \tag{3}$$

(We assume that $\alpha \log n$ is an integer.) So now let $Z_\alpha = Z_\alpha(SP)$ denote the cost of the minimum cost path from 1 to 2, i.e. two arbitrary fixed vertices. Then

Theorem 2.

$$Z_\alpha \sim \frac{\lambda \log n}{n} w.h.p. \tag{4}$$

where $0 < \alpha < 1/2$ and $\lambda = \lambda(\alpha)$ is the unique solution to

$$\frac{\lambda}{2} + \alpha \log\left(\frac{\lambda}{2\alpha}\right) + \alpha - 1 = 0.$$

Note that if we let $\alpha \nearrow 1/2$ then $\lambda \to 1$. This makes sense because the color constraint becomes essentially inactive, and Janson [10] proved that without the color constraint the minimum cost of a path from 1 to 2 is w.h.p. $\log n/n$. Similarly, if $\alpha \searrow 0$ then $\lambda \to 2$, reflecting the fact that we are only able to use blue edges.

At the present moment we do not have a proof that $\mathbb{E}Z_\alpha \sim \frac{\lambda \log n}{n}$, although we conjecture this to be true.

1.3 Assignments

For this problem each edge e of the complete bipartite graph is given a cost X_e independently distributed as exponential mean 1. Edges are then independently colored red or blue with probability 1/2. We seek a minimum cost perfect matching that contains at most αn red edges, where $0 < \alpha < 1/2$. (We assume that αn is an integer.) The cost $C(M)$ of a perfect matching M is given by $\sum_{e \in M} X_e$. In this context we prove

Theorem 3. *Let $Z_\alpha = Z_\alpha(PM)$ denote the minimum cost of a perfect matching M with $|E_R \cap M| \le \alpha n$, where $0 < \alpha < 1/2$. Let ψ be the solution to*

$$\psi^{-1} e^{-\psi/2} \int_{x=0}^{\psi/2} \frac{e^x - 1}{x} dx = \alpha.$$

Then

$$\mathbb{E}Z_\alpha \sim \frac{\pi^2}{3} + \int_{x=0}^{1} \frac{e^{-\psi x/2} \log(1-x)}{x} dx - \alpha\psi. \tag{5}$$

Furthermore, $Z_\alpha \sim \mathbb{E}Z_\alpha$ w.h.p.

We note that as $\alpha \nearrow 1/2$, we have $\psi \to 0$. Noting that $\int_{x=0}^{1} \frac{\log(1-x)}{x} \mathrm{d}x = -\pi^2/6$, we get $\mathbb{E}Z_\alpha \to \pi^2/6$. This is to be expected, since without the color constraint Aldous [1] proved that the expected minimum cost is $\pi^2/6$.

1.4 Asymmetric Traveling Salesperson Problem

For this problem each edge e of the complete digraph is given an independent exponential mean 1 cost X_e. Edges are then independently colored red or blue with probability 1/2. We seek a minimum cost tour T that contains at most αn red edges, where $\alpha < 1/2$. The cost of a tour T is given by $C(T) = \sum_{e \in T} X_e$. Karp [11] described a *patching algorithm* that w.h.p. finds a tour of a cost $1 + o(1)$ of the minimum. We use Karp's argument [11] and prove

Theorem 4. *Let $Z_\alpha(ATSP)$ denote the minimum cost of a Traveling Salesperson Tour T with $|E_R \cap T| \le \alpha n$. Then w.h.p.*

$$Z_\alpha(ATSP) \sim Z_\alpha(PM).$$

The above theorems are stated for specific distributions. It is easy to adapt them to more general distributions and this is done in Section 6.

2 Proof of Theorem 1

First we verify the value of f by evaluating the series in (1). Let W_0 be the principal branch of the Lambert W function, i.e. $W_0(z)$ is defined for $z \ge -1/e$, and returns the unique value $w \ge -1$ such that $we^w = z$.

Thus,

$$W_0(-xe^{-x}) = \begin{cases} -x, & x \le 1 \\ -y, & x > 1 \end{cases} \tag{6}$$

where $y < 1$ satisfies $ye^{-y} = xe^{-x}$. The Taylor series for W_0 is well known:

$$W_0(z) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}k^{k-1}}{k!} z^k.$$

Let

$$S(z) := \sum_{k=1}^{\infty} \frac{k^{k-2}}{k!} z^k \qquad \text{so} \qquad zS'(z) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} z^k = -W_0(-z).$$

Now since $S(0) = 0$ we have

$$S(z) = \int_0^z \frac{-W_0(-t)}{t} \mathrm{d}t = \int_0^{-W_0(-z)} \frac{w}{we^{-w}}(1-w)e^{-w}\mathrm{d}w = \int_0^{-W_0(-z)} (1-w)\,\mathrm{d}w = -W_0(-z) - \frac{1}{2}W_0(-z)^2$$

where we made the substitution $w = -W_0(-t)$, or $t = we^{-w}$. Now by the series definition of $f(x)$ in (1) and using (6), we get

$$f(x) = \frac{1}{x}\sum_{k=1}^{\infty} \frac{k^{k-2}}{k!}(xe^{-x})^k = \frac{1}{x}S(xe^{-x}) = \frac{1}{x}\left[-W_0(-xe^{-x}) - \frac{1}{2}W_0(-xe^{-x})^2\right] = \frac{1}{x} \cdot \begin{cases} x - \frac{1}{2}x^2, & x \le 1 \\ y - \frac{1}{2}y^2, & x > 1 \end{cases}$$

which matches the final expression in (1).

2.1 Dual problem

Let X denote the set of spanning trees of K_n. Our optimization problem is

$$\begin{aligned} \text{Minimise:} \quad & C(T) && (7) \\ \text{Subject to:} \quad & T \in X \\ & |E_R \cap T| \le \alpha n. && (8) \end{aligned}$$

Here we have a classical problem in combinatorial optimization, viz. find the minimum cost set that is a basis of one matroid (the cycle matroid of K_n) and independent in a second matroid (the partition matroid defined by (8)).

We analyse this through Lagrangian relaxation. Define the dual function $\phi(\theta)$ for $\theta \ge 0$ by

$$\phi(\theta) = -\theta\alpha n + L_n(\theta). \tag{9}$$

where $L_n(\theta) = \min_{T \in X}\{C_\theta(T)\}$ where $C_\theta(T) = \sum_{e \in T} C_\theta(e)$ and $C_\theta(e) = X_e + \theta 1_{e \in E_R}$,

Let θ^* be the value of θ maximizing $\phi(\theta)$ over all $\theta \ge 0$. It then follows from Edmonds' theorem [5] on the intersection of matroid polyhedra and Geoffrion's theorem on Lagrangian relaxation [8] that

$$\phi(\theta^*) = \max_{\theta \ge 0} \phi(\theta) = \min_{T \in X} \{C(T) : (8) \text{ holds}\},$$

To see this, observe that problem (7) is actually a linear program, since as observed in [8], ϕ finds an $\mathbf{x}$ that solves the linear program where $T \in X$ is replaced by $\mathbf{x} \in conv(X)$. And then [5] shows that in the case of

matroid intersection adding (8) does not create new fractional vertices i.e. the incidence vectors of trees with few enough red edges define the set of extreme points

We also know that w.h.p. the minimum cost of a spanning tree that only uses blue edges is asymptotic to $2\zeta(3) = 2.40411380632...$, see Beveridge, Frieze and McDiarmid [3]. So,

$$\zeta(3) \lesssim \phi(\theta^*) \lesssim 2\zeta(3). \tag{10}$$

2.2 Analysis of ϕ

Let the $U_e, e \in E(K_n)$ be a collection of independent and identically distributed (i.i.d.) uniform random variables on $[0,1]$ and let $B_e, e \in K_n$ be a collection of i.i.d. Bernoulli random variables with success probability $1/2$. For each $e \in E(K_n)$ set

$$W_{\theta,e} = U_e + \theta B_e \text{ and } L_n(\theta) = \min_{T \in X} \sum_{e \in T} W_{\theta,e}.$$

Then

$$\mathbb{E}\phi(\theta) = -\theta\alpha n + \mathbb{E}L_n(\theta).$$

Following Janson [9] we write

$$\mathbb{E}L_n(\theta) = \mathbb{E}\left(\sum_{e \in T_0} W_{\theta,e}\right) = \mathbb{E}\left(\sum_{e \in T_0} \int_{x=0}^{\infty} 1_{x < W_{\theta,e}} dx\right) = \int_{x=0}^{\infty} \mathbb{E}(\kappa(G_{n,\hat{p}(x)}) - 1)\mathrm{d}x,$$

where $\kappa(G)$ denotes the number of connected components of the graph $G = G_{n,\hat{p}(x)}$, and

$$\hat{p}(x) = \mathbb{P}(W_{\theta,e} \le x) = \frac{1}{2}\mathbb{P}(U_e \le x) + \frac{1}{2}\mathbb{P}(U_e \le x - \theta) = \frac{1}{2}\min\{x, 1\} + \frac{1}{2}\min\{(x-\theta)^+, 1\}.$$

Case 1.1. $\theta \ge K/n$ for some large constant K. The dual function $\phi(\theta) = L_n(\theta) - \theta\alpha n$ and w.h.p. $L_n(\theta) = O(1)$, independent of θ, given that we can just use blue edges for a tree. We have to maximise this over $\theta \ge 0$. This rules out $\theta \ge K/n$ for some constant $K > 0$, as we know that $\max \phi(\theta) \ge \zeta(3)$.

Case 1.2. $0 < \theta < K/n$. Then, using $\theta < 1$ we have

$$\hat{p}(x) = \begin{cases} x/2, & 0 < x < \theta, \\ x - \theta/2, & \theta < x < 1, \\ (1 + x - \theta)/2, & 1 < x < 1 + \theta, \\ 1, & x > 1 + \theta. \end{cases}$$

Thus, since $\hat{p}(x) = 1$ implies that $\kappa(G_{n,\hat{p}(x)}) = 1$,

$$\begin{aligned}
\mathbb{E}L_n(\theta) &= \int_{x=0}^{1+\theta} \mathbb{E}(\kappa(G_{n,\hat{p}(x)}) - 1)\mathrm{d}x \\
&= \int_0^{\theta} \mathbb{E}(\kappa(G_{n,x/2}))\mathrm{d}x + \int_{\theta}^{1} \mathbb{E}(\kappa(G_{n,x-\theta/2}))\mathrm{d}x + \int_1^{1+\theta} \mathbb{E}(\kappa(G_{n,(1+x-\theta)/2}))\mathrm{d}x - (1+\theta) \\
&= 2\int_0^{\theta/2} \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}x + \int_{\theta/2}^{1-\theta/2} \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}x + 2\int_{1-\theta/2}^{1} \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}x - 1 + O(1/n), \\
&= \int_0^{\theta/2} \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}x + \int_0^{1} \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}x - 1 + O(1/n) \\
&= \int_0^{\theta/2} \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}x + \zeta(3) + O(1/n), \qquad \text{uniformly for all } \theta < K/n.
\end{aligned}$$

Explanation: on the third line each of the three integrals is the same as one of the integrals in the previous line after making an appropriate substitution. On the fourth line we have used the fact that for $x \geq 1 - O(1/n)$ we have $\mathbb{E}(\kappa(G_{n,x})) = 1 + o(1)$. On the final line we used the fact that $\int_0^1 \mathbb{E}(\kappa(G_{n,x}))\mathrm{d}p - 1 = \zeta(3) + O(1/n)$, due to Janson [9].

Claim 1. $\mathbb{E}(\kappa(G_{n,p})) = nf(np) + o(n)$ *uniformly for* $p < K/n$.

Proof of claim. Now $\mathbb{E}(\kappa(G_{n,p})) = \mathbb{E}(A) + \mathbb{E}(B) + \mathbb{E}(C)$ where A is the number of tree components on at most $\log^2 n$ vertices, and B is the nontrees on at most $\log^2 n$ vertices, and C is all components on more than $\log^2 n$ vertices. B and C will be small. Indeed, deterministically we have $C \leq n/\log^2 n$. For B we have

$$\mathbb{E}(B) \leq \sum_{k=1}^{\log^2 n} \binom{n}{k} k^{k-2} \binom{k}{2} p^k (1-p)^{k(n-k)} \leq \sum_{k=1}^{\log^2 n} \frac{(knpe^{-np})^k}{k!} (1-p)^{-\log^4 n} = O(\log^2 n),$$

where the final bound is because $ze^{-z} \leq 1/e$ for all $z > 0$, $k! > (k/e)^k$, and $(1-p)^{-\log^4 n} = 1 + o(1)$. Now for A we have

$$\begin{aligned}
\mathbb{E}(A) &= \sum_{k=1}^{\log^2 n} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{k}{2}-k+1} \\
&= \left(1 + O\left(\frac{\log^4 n}{n}\right)\right) \frac{1}{p} \sum_{k=1}^{\log^2 n} \frac{k^{k-2}}{k!} (np)^k e^{-knp} = nf(np) + o(n).
\end{aligned}$$

□

In the next subsection we will show that $\mathbb{E}L_n(\theta)$ is highly concentrated. Thus we have, on making the substitution $u = nx$ in the integral,

$$\phi(\theta) = -\theta\alpha n + \int_0^{n\theta/2} f(u)\mathrm{d}u + \zeta(3) + o(1)$$

w.h.p. Now differentiating the above w.r.t. θ and setting it equal to 0, we see that $\theta^* = 2\psi/n$ where $f(\psi) \sim 2\alpha$. Thus w.h.p.

$$\phi(\theta^*) = -2\alpha\psi + \int_0^{\psi} f(u)\mathrm{d}u + \zeta(3) + o(1),$$

and Theorem 1 follows.

2.3 Concentration

The goal of this section is to prove the following lemma.

Lemma 5. *For a fixed* $0 \le \theta \le K/n$,

$$\mathbb{P}(|L_n(\theta) - \mathbb{E}(L_n(\theta))| \ge n^{-1/5}) = o(n^{-100}). \tag{11}$$

Proof. Recall that $\phi(\theta) = L_n(\theta) - \theta\alpha n$ (as defined in (9)). We will use Theorem 8.1.1 of Talagrand [13]. Let $0 < X_i \le \xi, i = 1, 2, \ldots, N$ be a sequence of independent positive random variables and let $\mathcal{F}$ be a family of N-tuples $\boldsymbol{\alpha} = \alpha_i, i = 1, 2, \ldots, N$. Let

$$Z = \inf_{\boldsymbol{\alpha} \in \mathcal{F}} \sum_{i=1}^{N} \alpha_i X_i. \tag{12}$$

Let $\sigma = \sup_{\boldsymbol{\alpha} \in \mathcal{F}} \left(\sum_{i=1}^{N} \alpha_i^2 \right)^{1/2}$ and let μ be a median of Z. Then, for all $u > 0$,

$$\mathbb{P}(|Z - \mu| \ge u) \le 4 \exp\left\{ -\frac{u^2}{4\sigma^2} \right\}. \tag{13}$$

We need the following claim: let T_θ denote the minimum cost spanning tree with costs C_θ.

Claim 2. *Let* $0 \le \theta \le K/n$. *Then* $\mathbb{P}(\max_{e \in T_\theta} W_{e,\theta} \ge n^{-9/10}) = o(n^{-200})$.

Proof of claim Let $L = n^{-9/10}$. Just consider running Kruskal's greedy algorithm. The probability that the graph is not connected by the time we reach edges of cost $W_{\theta,e} \le L/2$ is certainly $o(n^{-200})$. ($U_e \le p = L/2 - \theta$ implies that $W_{\theta,e} \le L/2$ and $\mathbb{P}(G_{n,p}$ is not connected) is at most $\sum_{k=1}^{n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} = \sum_{k=1}^{n/2} O(n^{-1}(npe^{-np})^k) = o(n^{-200})$.)
End of proof of claim

Let $X_e = n^{9/10} \min\left\{ C_\theta(e), n^{-9/10} \right\}$ for $e \in E(K_n)$. The claim implies that if we compute the minimum spanning tree with costs X_e, then with probability $1 - o(n^{-200})$ we get the same spanning tree as if we used costs $C_\theta(e)$, but with cost inflated by a factor $n^{9/10}$. To apply (13) we let $N = \binom{n}{2}$ and we let $\boldsymbol{\alpha}$ be the 0/1 incidence vectors of the spanning trees of K_n. Thus in this case, Z of (12) is the cost of the minimum spanning tree using costs X_e. Now $\sigma \le n^{1/2}$ and $u = n^{3/5}$ so that from (13) we have

$$\mathbb{P}(|Z - \mu| \ge n^{3/5}) \le 4 \exp\left\{ -\frac{n^{1/5}}{4} \right\} = o(n^{-200}).$$

It follows that

$$\mathbb{P}(|L_n(\theta) - \widehat{\mu}| \ge n^{-3/10}) = o(n^{-200}) \tag{14}$$

where $\widehat{\mu} = \mu n^{-9/10}$. It follows immediately that

$$\mathbb{E}L_n(\theta) \ge \widehat{\mu} - 2n^{-3/10}. \tag{15}$$

On the other hand

$$\mathbb{E}L_n(\theta) \le \widehat{\mu} + n^{-3/10} + n^2 \mathbb{P}(L_n(\theta) \ge \widehat{\mu} + n^{-3/10}) \le \widehat{\mu} + 2n^{-3/10}. \tag{16}$$

So, from (14), (15), (16) we get $|L_n(\theta) - EL_n(\theta)| \le 4n^{-3/10}$ with probability $1 - o(n^{-100})$. □

The fixed-θ estimate above implies a uniform estimate on $[0, K/n]$. Indeed, $L_n(\theta)$ is $(n-1)$-Lipschitz in θ, and $-\theta\alpha n$ is αn-Lipschitz. Hence $\phi(\theta)$ is at most $2n$-Lipschitz. Take a grid of mesh n^{-4} in $[0, K/n]$, containing $O(n^3)$ points. The fixed-θ concentration bound, union-bounded over the grid, gives concentration at every grid point with probability $1 - o(n^{-90})$. Lipschitz continuity then gives $\sup_{0\le\theta\le K/n} |\phi(\theta) - \mathbb{E}\phi(\theta)| = o(1)$ w.h.p. Therefore $\max_\theta \phi(\theta) = \max_\theta \mathbb{E}\phi(\theta) + o(1)$ w.h.p. Since the limiting value in Theorem 1 is bounded away from zero, this absolute $o(1)$ concentration implies $Z_\alpha \sim \mathbb{E}Z_\alpha$ w.h.p.

3 Proof of Theorem 2

The proof here is written for the undirected case K_n, but the same argument will work for the complete digraph $\vec{K}_n$ if we use outgoing edges.

3.1 A lower bound

We prove that for every fixed $\eta < \lambda(\alpha)$,

$$\mathbb{P}\left(Z_\alpha \le \eta\frac{\log n}{n}\right) \longrightarrow 0.$$

Let

$$t_n := \eta\frac{\log n}{n},$$

and let $N_n(\eta)$ be the number of paths from 1 to 2 with total weight at most t_n and at most L red edges. A fixed k-edge path has weight Γ_k where

$$\mathbb{P}\big(\Gamma_k \le t_n\big) = \int_0^{t_n} \frac{x^{k-1}e^{-x}}{(k-1)!}\,\mathrm{d}x \le \frac{t_n^k}{k!},$$

where the inequality follows from $e^{-x} \le 1$. There are at most n^{k-1} paths of length k from 1 to 2. For such a path, the red-edge count has distribution $Bin(k, 1/2)$, independently of the weights. Therefore

$$\begin{aligned}
\mathbb{E}N_n(\eta) &\le \sum_{k=1}^{n-1} n^{k-1}\frac{t_n^k}{k!}\mathbb{P}\big(Bin(k,1/2) \le L\big)\\
&= \frac{1}{n}\sum_{k=0}^{\infty}\frac{(\eta\log n)^k}{k!}\sum_{\ell=0}^{L}\binom{k}{\ell}2^{-k}\\
&= \frac{1}{n}\sum_{\ell=0}^{L}\frac{1}{\ell!}\left(\frac{\eta\log n}{2}\right)^\ell\sum_{k=\ell}^{\infty}\frac{(\eta\log n)^{k-\ell}}{2^{k-\ell}(k-\ell)!}\\
&= \frac{1}{n^{1-\eta/2}}\sum_{\ell=0}^{L}\frac{1}{\ell!}\left(\frac{\eta\log n}{2}\right)^\ell. \qquad (17)
\end{aligned}$$

If $\eta \le 2\alpha < 1$ then (17) gives

$$\mathbb{E}N_n(\eta) \le n^{\eta-1} \to 0.$$

It remains to consider $\eta \ge 2\alpha$. Now the sum in (17) is dominated by its last term, so

$$\frac{1}{n^{1-\eta/2}}\sum_{\ell=0}^{L}\frac{1}{\ell!}\left(\frac{\eta\log n}{2}\right)^\ell = O\left(\frac{1}{n^{1-\eta/2}}\cdot\frac{1}{L!}\left(\frac{\eta\log n}{2}\right)^L\right) = O(n^{-1+\eta/2+\alpha(1+\log\eta-\log 2\alpha)}) = o(1),$$

since $\eta < \lambda(\alpha)$.

3.2 An upper bound

The idea here is to grow shortest path trees from 1 and 2, up to certain depth. Taking account of the coloring, we argue that each tree contains a large number of vertices whose path from the root has small cost and uses fewer than $L/2$ red edges. Then it is simple to argue that w.h.p. we can join two such vertices to make a short path from 1 to 2.

Consider the analysis of Janson [10] in the uncolored case. Recall Dijkstra's algorithm DA. After k iterations there is a rooted *shortest path* tree T_k with root $s = 1$ such that if v is a vertex of T_k then the tree path from s to v is a shortest path. Let $d(v)$ be its length. For $x \notin T_k$ let $d(x)$ be the minimum length of a path P that goes from s to v to x where $v \in T_k$ and the sub-path of P that goes to v is the tree path from s to v. If $d(y) = \min\{d(x) : x \notin T_k\}$ then $d(y)$ is the length of a shortest path from s to y and y can be added to the tree.

Suppose that vertices are added to the tree in the order $v_1 = s, v_2, \ldots, v_n$ and that $Y_j = dist(v_1, v_j)$ for $j = 1, 2, \ldots, n$. It follows from the memoryless property of the exponential distribution[1]

$$Y_{k+1} = \min_{\substack{i=1,2,\ldots,k \\ v \neq v_1,\ldots,v_k}} [Y_i + X_{v_i,v}] = Y_k + E_k$$

where E_k, being the minimum of $k(n-k)$ independent exponentials, i.e. exponential with mean $\frac{1}{k(n-k)}$ and independent of Y_k. This is because X_{v_i,v_j} is distributed as an independent exponential Z conditioned on $Z \geq Y_k - Y_i$. Hence

$$\mathbb{E}Y_n = \sum_{k=1}^{n-1} \frac{1}{k(n-k)} = \frac{1}{n} \sum_{k=1}^{n-1} \left(\frac{1}{k} + \frac{1}{n-k} \right) = \frac{2}{n} \sum_{k=1}^{n-1} \frac{1}{k} = \frac{2 \log n}{n} + O(n^{-1}). \tag{18}$$

Now consider what changes when we color the edges. We apply the Dijkstra's algorithm (DA) ignoring the colors and produce a sequence of trees $T_1, T_2, \ldots, T_n$ where T_k has exactly k vertices. A vertex v of T_n is *valid* if the path from 1 to v contains at most L red edges after the edge colors are revealed.

Suppose that vertices are added to the tree in the order $v_1 = s, v_2, \ldots, v_n$. Then if $i \leq k < j$, the edge cost X_{v_i,v_j} is distributed as an independent exponential $Z_{i,j}$ conditioned on $Z_{i,j} \geq d(v_k) - d(v_i)$. It follows from the memoryless property of the exponential distribution that $Z_{i,j} = d(v_k) - d(v_i) + E_{i,j}$ where the $E_{i,j}$ are independent exponential mean 1. Then,

$$\min_{1 \leq i \leq k < j} (d(v_i) + Z_{i,j}) = d_k + \min_{1 \leq i \leq k < j} E_{i,j}.$$

It follows that when we add a new vertex it is connected to a random vertex in the existing tree.

3.3 A colored Yule process

We model the growth of the trees $T_1, T_2, \ldots, T_n$ as a *Yule process*. Consider a continuous-time branching process started from one root at time 0. Each particle lives forever and gives birth at rate 1. At each birth,

[1] $\mathbb{P}(Z \geq A + t \mid Z \geq A) = \mathbb{P}(Z \geq t)$ for an exponential random variable Z

the child is joined to its parent by an independently colored edge, red with probability 1/2 and blue with probability 1/2. In algorithm DA, the particles are the vertices, the root is 1 and when a child is born, it is born to a random parent. Denote the Yule tree constructed by time t as $\mathcal{Y}(t)$.

In both processes when a vertex is added to a tree, it is connected to a uniform random neighbor. (Such trees are called *Random recursive trees* and have been well studied.) We are only going to use valid paths of $\mathcal{Y}(t)$. If v is the kth vertex added to $\mathcal{Y}(t)$ and the path P_k to v is valid then the expected corresponding Dijkstra cost of P_k is

$$\frac{1}{n}\sum_{i=1}^{k}\left(\frac{1}{i}+\frac{1}{n-i}\right)\sim\frac{\log k}{n}+O\left(\frac{1}{n}\right), \tag{19}$$

if $k=o(n)$.

Meanwhile the kth birth of $\mathcal{Y}$ occurs at expected time $\sim\sum_{i=1}^{k}\frac{1}{i}\sim\log k$. This is because the time to the next birth when there are i vertices in $\mathcal{Y}$ is the minimum of i independent exponentials with mean 1. Meanwhile, having explored k vertices, the next increment in cost has distribution $EXP(k(n-k))$, so for $k=o(n)$ the scaled process $nd(v)$ has increments close to those of a rate-k Yule process. The parent of the new vertex is uniform among the k existing vertices, and the edge color is independent of the costs and of the past. Thus the rooted colored tree structure agrees with a colored random recursive tree up to $o(n)$ explored vertices, while distances are Yule times divided by n, up to a $1+o(1)$ factor.

We will prove the following lemma in Section 3.4 below;

Lemma 6. *Fix $a,\eta\in(0,1)$, $\gamma\in(0,a/2)$ and $s<h(a,\gamma)$ where*

$$h(a,\gamma)=\frac{a}{2}+\gamma+\gamma\log\left(\frac{a}{2\gamma}\right).$$

In the complete graph $K_{n-o(n)}$ with independent $EXP(1)$ edge costs, run algorithm DA until $N=n^{1-\eta}$ vertices have been discovered. Then w.h.p. there are at least N^s vertices $v\in T_N$ such that the path in T_N from 1 to v has total cost at most $\frac{(a+o(1))\log N}{n}$ and uses at most $\gamma\log N$ red edges.

Assume the truth of Lemma 6 with $\eta>0$ is sufficiently small. We choose $a=\lambda/2+\varepsilon,\gamma=\alpha/2$ where $\varepsilon>0$ is sufficiently small. We argue below that for some $c_1>0$, we have

$$h(a,\gamma)\geq\frac{1}{2}+c_1\varepsilon+O(\varepsilon^2). \tag{20}$$

Given (20), choose s,η so that $1/2<s(1-2\eta)<s<h(a,\gamma)$ and fix a set W of size $n^{3/2-s}$. Apply Lemma 6 and run DA on $K_n\setminus W$ from vertex 1 until we have explored $N=n^{1-\eta}$ vertices and found $N^s=n^{(1-\eta)s}$ valid vertices V_1 as promised by Lemma 6. Now do the same from vertex 2, running DA as if on $K_{n-N}\setminus W$, avoiding the $N=o(n)$ vertices examined for vertex 1 until we obtain N^s valid vertices V_2. Note that the costs of the edges between V_1,V_2 and W are unconditioned. If there is a path $P=(w_1\in V_1,w\in W,w_2\in V_2)$ using blue edges of cost at most $1/n$ then concatenating the path from 1 to w_1 in V_1 and the path P and the path from w_2 to 2 in V_2 gives us (after possibly removing cycles) a path of cost at most $(\lambda+2\varepsilon+o(1))\log n/n$ containing at most $\alpha\log n$ red edges. Now the probability that there is no such path P can be bounded

$$(1-(1-(1-\tfrac{1}{2}(1-e^{-1/n}))^{n^{(1-\eta)s}})^2)^{n^{3/2-s}}=e^{-\Omega(n^{(1-2\eta)s-1/2})}=o(1).$$

Since ε is arbitrary, Theorem 2 follows.

Verification of (20) We first observe that from the definition of α, we have

$$h\left(\frac{\lambda}{2}, \frac{\alpha}{2}\right) = \frac{1}{2}\left(\frac{\lambda}{2} + \alpha + \alpha \log\left(\frac{\lambda}{2\alpha}\right)\right) = \frac{1}{2}.$$

Then we observe that

$$\nabla h\left(\frac{\lambda}{2}, \frac{\alpha}{2}\right) = \left(\frac{1}{2} + \frac{\alpha}{\lambda}, \log\left(\frac{\lambda}{2\alpha}\right)\right) > 0. \tag{21}$$

since $\lambda > 2\alpha$. (If $\lambda \leq 2\alpha$ then its definition implies that $\lambda/2 \geq 1 - \alpha > \alpha$, contradiction.)

3.4 Proof of Lemma 6: Structure of $\mathcal{Y}(t)$

We analyze $\mathcal{Y}(t)$, proving the analog of Lemma 6 for the Yule process, which we now state.

Lemma 7. *Fix $a \in (0,1)$, $\gamma \in (0, a/2)$ and $s < h(a,\gamma)$ where*

$$h(a,\gamma) = \frac{a}{2} + \gamma + \gamma \log\left(\frac{a}{2\gamma}\right).$$

In the Yule process, the following holds w.h.p. There are at least N^s particles v such that the path from the root to v appears by time $(a + o(1)) \log N$ and uses at most $\gamma \log N$ red edges.

Lemma 7 implies Lemma 6. Indeed, we can couple the DA exploration process up to step N with the Yule process, where to calculate the DA costs from the Yule costs we divide by $n' = n - o(n)$. In the DA process, at step k the next vertex arrival is distributed as $EXP(k(n'-k))$ and for $k \leq N = o(n)$ we have $k(n'-k) \sim kn$.

For $\theta > 0$ define

$$S_\theta(t) := \sum_{u \in \mathcal{Y}(t)} e^{-\theta R_u}, \qquad \psi(\theta) := \frac{1 + e^{-\theta}}{2}, \qquad M_\theta(t) := e^{-\psi(\theta)t} S_\theta(t),$$

where R_u denotes the number of red edges in the path from the root to u in $\mathcal{Y}(t)$.

Lemma 8. *For every fixed $\theta > 0$, $(M_\theta(t))_{t \geq 0}$ is a nonnegative martingale, bounded in L^2, and it converges almost surely and in L^2 to a finite limit $M_\theta(\infty)$ satisfying*

$$\mathbb{P}(M_\theta(\infty) > 0) = 1.$$

Consequently, for every fixed $a > 0$,

$$\frac{\log M_\theta(a \log N)}{\log N} \xrightarrow{\mathbb{P}} 0.$$

Proof. The martingale property follows directly from the generator $\mathcal{L}$ of the Yule branching process. $\mathcal{L}$ acts on a function of the current tree by computing its instantaneous expected rate of change. If a particle u has red-depth $R_u = r$, then a blue child is born at rate $\rho = 1/2$ and contributes $e^{-\theta r}$ to S_θ, while a red child is born at rate ρ and contributes $e^{-\theta(r+1)}$. Hence

$$\mathcal{L} S_\theta = \rho \sum_{u \in \mathcal{Y}(t)} \left(e^{-\theta R_u} + e^{-\theta(R_u+1)}\right) = \rho(1 + e^{-\theta}) S_\theta = \psi(\theta) S_\theta.$$

Therefore, the instantaneous expected rate of change for $F(t, \mathcal{Y}(t)) = M_\theta(t)$ is

$$(\partial_t + \mathcal{L})F(t, \mathcal{Y}(t)) = -\psi(\theta)e^{-\psi(\theta)t}S_\theta + e^{-\psi(\theta)t}\mathcal{L}S_\theta = 0.$$

(Here ∂_t is the deterministic contribution due to the factor $e^{-\psi(\theta)t}$ and $\mathcal{L}$ is due to the stochastic drift.)

We use Dynkin's formula [4] in its time-dependent form. If X_t is a continuous-time Markov process with generator $\mathcal{L}$ and F is a suitable function, then

$$\mathbb{E}\left(F(t, X_t) \mid \mathcal{F}_s\right) = F(s, X_s) + \mathbb{E}\left(\int_s^t (\partial_r + \mathcal{L})F(r, X_r)\, dr \mid \mathcal{F}_s\right).$$

In particular, if $(\partial_t + \mathcal{L})F(t, X_t) = 0$, then $F(t, X_t) = M_\theta(t)$ is a martingale.

For a second moment calculation, observe that a birth increases the quantity $S_\theta(t)^2$ by $2S_\theta(t)Y + Y^2$. We have

$$\mathbb{E}(Y^2) = \rho \sum_{u \in \mathcal{Y}(t)} (e^{-2\theta R_u} + e^{-2\theta(R_u+1)}) = \rho S_{2\theta}(t) + \rho e^{-2\theta} S_{2\theta}(t) = \psi(2\theta)S_{2\theta}(t).$$

Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(S_\theta(t)^2) = 2\psi(\theta)\mathbb{E}(S_\theta(t)^2) + \psi(2\theta)\mathbb{E}S_{2\theta}(t) = 2\psi(\theta)\mathbb{E}(S_\theta(t)^2) + \psi(2\theta)e^{\psi(2\theta)t}.$$

Because $\psi(2\theta) < 2\psi(\theta)$ for $\theta > 0$, integrating gives us

$$\mathbb{E}(S_\theta(t)^2) = e^{2\psi(\theta)t}\left(1 + \psi(2\theta)\int_0^t e^{(\psi(2\theta)-2\psi(\theta))s}\mathrm{d}s\right) \le e^{2\psi(\theta)t}\left(1 + \frac{\psi(2\theta)}{2\psi(\theta) - \psi(2\theta)}\right).$$

So,

$$\mathbb{E}(M_\theta(t)^2) = e^{-2\psi(\theta)t}\mathbb{E}(S_\theta(t)^2) \le 1 + \frac{\psi(2\theta)}{2\psi(\theta) - \psi(2\theta)} < \infty$$

uniformly in t. Thus, by the Martingale Convergence Theorem, $M_\theta(t)$ converges almost surely and in L^2 to a limit $M_\theta(\infty)$. Note that $\mathbb{E}M_\theta(\infty) = \mathbb{E}M_\theta(0) = 1$.

It remains only to check that this limit is almost surely positive. Fix $T > 0$. By the branching property,

$$M_\theta(\infty) = \sum_{u \in \mathcal{Y}(T)} e^{-\theta R_u - \psi(\theta)T} M_\theta^{(u)}(\infty),$$

where, conditional on the process up to time T, the random variables $M_\theta^{(u)}(\infty)$ are independent copies of $M_\theta(\infty)$. Let $q := \mathbb{P}(M_\theta(\infty) = 0)$. Since all coefficients in the last sum are positive,

$$q = \mathbb{E}(q^{|\mathcal{Y}(T)|}).$$

The Yule population satisfies $|\mathcal{Y}(T)| \ge 1$ almost surely and $\mathbb{P}(|\mathcal{Y}(T)| > 1) > 0$. Hence, for $0 < q < 1$, one has $\mathbb{E}(q^{|\mathcal{Y}(T)|}) < q$. Also $q = 1$ is impossible because $\mathbb{E}M_\theta(\infty) = 1$. Therefore $q = 0$.

The logarithmic statement follows immediately from almost sure convergence to a finite positive limit. □

Given the above lemma, we can give the proof of Lemma 7.

Since $\partial h(a, \gamma)/\partial\gamma = \log(a/2\gamma) > 0$, we may choose $\bar{\gamma} < \gamma$ such that $s < h(a, \bar{\gamma})$. Choose $\delta > 0$ such that $0 < \bar{\gamma} - \delta < \bar{\gamma} + \delta < \gamma$. Then with δ sufficiently small and $\theta = \log(a/2\bar{\gamma}) > 0$, we have $a\psi(\theta) + \theta(\bar{\gamma} - \delta) > s$. This is because

$$a\psi(\theta) + \theta\bar{\gamma} = \frac{a}{2} + \bar{\gamma} + \bar{\gamma}\log\left(\frac{a}{2\bar{\gamma}}\right) = h(a, \bar{\gamma}).$$

Let $t = a \log N$, then for any $u > 0$,

$$\sum_{R_v < (\bar{\gamma}-\delta)\log N} e^{-\theta R_v} \le e^{u(\bar{\gamma}-\delta)\log N} \sum_{v \in \mathcal{Y}(t)} e^{-(\theta+u)R_v} = \exp\left\{(u(\bar{\gamma}-\delta) + a\psi(\theta+u))\log N\right\} M_{\theta+u}(t).$$

The derivative with respect to u at $u = 0$ of $u(\bar{\gamma}-\delta) + a\psi(\theta+u) - a\psi(\theta)$ is $\bar{\gamma} - \delta - ae^{-\theta}/2 = \bar{\gamma} - \delta - \bar{\gamma} = -\delta$. Thus, for some small fixed $u > 0$ and some $\kappa > 0$,

$$u(\bar{\gamma}-\delta) + a\psi(\theta+u) \le a\psi(\theta) - \kappa.$$

By Lemma 8, $M_{\theta+u}(t) \le e^{\kappa \log N/2}$ w.h.p. Therefore w.h.p.

$$\sum_{R_v < (\bar{\gamma}-\delta)\log N} e^{-\theta R_v} \le \exp((a\psi(\theta) - \kappa/2)\log N)\,.$$

Analogously, choose $0 < u < \theta$. Then

$$\sum_{R_v > (\bar{\gamma}+\delta)\log N} e^{-\theta R_v} \le e^{-u(\bar{\gamma}+\delta)\log N} \sum_{v} e^{-(\theta-u)R_v} = \exp\left\{(-u(\bar{\gamma}+\delta) + a\psi(\theta-u))\log N\right\} M_{\theta-u}(t).$$

The derivative with respect to u at $u = 0$ of $-u(\bar{\gamma}+\delta) + a\psi(\theta-u) - a\psi(\theta)$ is $-\bar{\gamma} - \delta + ae^{-\theta}/2 = -\delta$. So, after possibly decreasing u, w.h.p.,

$$\sum_{R_v > (\bar{\gamma}+\delta)\log N} e^{-\theta R_v} \le \exp\left\{(a\psi(\theta) - \kappa/2)\log N\right\}.$$

On the other hand, Lemma 8 gives that w.h.p.,

$$\sum_{v \in \mathcal{Y}(t)} e^{-\theta R_v} = e^{\psi(\theta)t} M_\theta(t) \ge \exp\left\{(a\psi(\theta) - o(1))\log N\right\}.$$

Therefore w.h.p.,

$$\sum_{R_v \in [\bar{\gamma}\pm\delta]\log N} e^{-\theta R_v} \ge \exp\left\{(a\psi(\theta) - o(1))\log N\right\}. \tag{22}$$

Consequently, the number of particles in the sum on the LHS of (22) must be at least

$$\exp\left\{(a\psi(\theta) + \theta(\bar{\gamma}-\delta) - o(1))\log N\right\} \ge N^s.$$

Lemma 7 now follows since $\bar{\gamma} + \delta \le \gamma$.

4 Proof of Theorem 3

In this section, we proceed as for spanning trees and consider a dual problem. Thus let X denote the set of perfect matchings of $K_{n,n}$. Our optimization problem is to compute

$$\begin{aligned} C^* = \text{Minimum:} \quad & C(M) && (23)\\ \text{Subject to:} \quad & M \in X \\ & |E_R \cap M| \le \alpha n. && (24) \end{aligned}$$

We analyse this through Lagrangian relaxation. Define the dual function $\phi(\theta)$ for $\theta \geq 0$ by

$$\phi(\theta) = -\theta\alpha n + \min_{M \in X} \sum_{e \in M} (X_e + \theta 1_{e \in E_R}). \tag{25}$$

Let $Z_e, e \in E(K_{n,n})$ be i.i.d. exponential mean 1, and let $\widehat{Z}_e$ be i.i.d. Bernoulli 1/2. For each $e \in E(K_{n,n})$ set

$$W_{\theta,e} = Z_e + \theta \widehat{Z}_e \text{ and } L_n(\theta) = \min_{M \in X} W_\theta(M) \text{ where } W_\theta(M) = \sum_{e \in M} W_{\theta,e}. \tag{26}$$

Then

$$\mathbb{E}\phi(\theta) = -\theta\alpha n + \mathbb{E}L_n(\theta).$$

Consider the complete bipartite graph $K_{n,n}$, with bipartition (A, B), where $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$, and with edge weights which are independent copies of $W_{\theta,e}$. We assume that $a_1, a_2, \ldots, a_n$ is a random permutation of A. So $A_r = \{a_1, a_2, \ldots, a_r\}$ is a random r-subset of A.

Let $EXP(\lambda)$ denote an exponential random variable of rate λ i.e. $\mathbb{P}(EXP(\lambda) \geq x) = e^{-\lambda x}$. We add a special vertex b^* to B, with edges to all n vertices of A. Each edge adjacent to b^* is assigned an $EXP(\lambda)$ weight independently, $\lambda > 0$.

For $r \geq 1$ we let M_r be the minimum weight matching of A_r into B and M_r^* be the minimum weight matching of A_r into $B^* = B \cup \{b^*\}$. (As $\lambda \to 0$ it becomes increasingly unlikely that any of the extra edges are actually used in the minimum weight matching.) We denote this matching by M_r^* and we let B_r^* denote the corresponding set of vertices of B^* that are covered by M_r^*. We let $C(n, r)$ denote the weight of M_r.

Define $P(n, r)$ as the normalized probability that b^* participates in M_r^*, i.e.

$$P(n, r) = \lim_{\lambda \to 0^+} \frac{\mathbb{P}(b^* \in B_r^*)}{\lambda}. \tag{27}$$

Its importance lies in the following lemma:

Lemma 9.

$$\mathbb{E}(C(n, r) - C(n, r-1)) = \frac{P(n, r)}{r}. \tag{28}$$

Proof. Choose i randomly from $[r]$ and let $\widehat{B}_i \subseteq B_r$ be the B-vertices in the minimum weight matching of $(A_r \setminus \{a_i\})$ into B^*. Let $X = C(n, r)$ and let $Y = C(n, r-1)$. Let w_i be the weight of the edge (a_i, b^*), and let I_i denote the indicator variable for the event that the minimum weight of an A_r matching that contains this edge is smaller than the minimum weight of an A_r matching that does not use b^*. We claim that $\mathbb{P}(I_i) = \mathbb{P}(Y_i + w_i < X) + O(\lambda^2)$, where Y_i is the minimum weight of a matching from $A_r \setminus \{a_i\}$ to B. Indeed, if $(a_i, b^*) \in M_r^*$ then $w_i < X - Y_i$. Conversely, if $w_i < X - Y_i$ and no other edge from b^* has weight smaller than $X - Y_i$, then $(a_i, b^*) \in M_r^*$, and when $\lambda \to 0$, the probability that there are two distinct edges from b^* of weight smaller than $X - Y_i$ is of order $O(\lambda^2)$. Indeed, let $\mathcal{F}$ denote the existence of two distinct edges from b^* of weight smaller than X and let $\mathcal{F}_{i,j}$ denote the event that (a_i, b^*) and (a_j, b^*) both have weight smaller than X.

Then,

$$\mathbb{P}(\mathcal{F}) \leq n^2 \mathbb{E}_X(\max_{i,j} \mathbb{P}(\mathcal{F}_{i,j} \mid X)) = n^2 \mathbb{E}((1 - e^{-\lambda X})^2) \leq n^2 \lambda^2 \mathbb{E}(X^2), \tag{29}$$

and since $\mathbb{E}(X^2)$ is finite and independent of λ, this is $O(\lambda^2)$.

Note that Y and Y_i have the same distribution. They are both equal to the minimum weight of a matching of a random $(r-1)$-set of A into B. As a consequence, $\mathbb{E}(Y) = \mathbb{E}(Y_i) = \frac{1}{r}\sum_{j\in A_r}\mathbb{E}(Y_j)$. Since w_i is $EXP(\lambda)$ distributed, as $\lambda \to 0$ we have from (29) that

$$P(n,r) = \lim_{\lambda\to 0}\left(\frac{1}{\lambda}\sum_{j\in A_r}\mathbb{P}(w_j < X - Y_j) + O(\lambda)\right) =$$

$$\lim_{\lambda\to 0}\mathbb{E}\left(\frac{1}{\lambda}\sum_{j\in A_r}\left(1 - e^{-\lambda(X-Y_j)}\right)\right) = \sum_{j\in A_r}\mathbb{E}(X - Y_i) = r\mathbb{E}(X-Y).$$

□

We now proceed to estimate $P(n,r)$. Fix r and assume that $b^* \notin B^*_{r-1}$. Note that

$$M^*_r \text{ is obtained from } M^*_{r-1} \text{ via a single augmenting path with } a_r \text{ as an endpoint. In particular, } B^*_{r-1} \subseteq B^*_r. \tag{30}$$

Indeed, clearly the symmetric difference $M^*_{r-1}\Delta M^*_r$ contains some M^*_{r-1}-augmenting path with a_r as an endpoint, since M^*_r covers a_r and M^*_{r-1} does not. Since both matchings cover $a_1,\ldots,a_{r-1}$, any other component of $M^*_{r-1}\Delta M^*_r$ must be either a path or cycle with an even number of edges. But both matchings are supposed to minimize cost, so they could not disagree on such a component. Thus there are no such components and (30) is proved.

Suppose that M^*_r is obtained from M^*_{r-1} by finding an augmenting path $P = (a_r,\ldots,a_\sigma,b_\tau)$ from a_r to $B\setminus B_{r-1}$ of minimum additional weight. We condition on (i) σ, (ii) the costs of all edges other than $(a_\sigma,b_j), b_j \in B\setminus B_{r-1}$ and
(iii) $\min\{w(a_\sigma,b_j) : b_j \in B\setminus B_{r-1}\}$. With this conditioning $M_{r-1} = M^*_{r-1}$ will be fixed and so will $P' = (a_r,\ldots,a_\sigma)$. Let $m = n-r+2$ and $X_1,X_2,\ldots,X_{m-1}$ be independent copies of $W_{\theta,e}$ and let X_m be exponential with rate λ. We compute the probability that X_m is the smallest of them if we condition on the value of $\min\{X_1,X_2,\ldots,X_m\}$. Indeed,

Claim 3. *For any $\gamma > 0$, we have*

$$\mathbb{P}(X_m = \min\{X_1,\ldots,X_m\} \mid \min\{X_1,\ldots,X_m\} = \gamma) = \begin{cases} \frac{\lambda}{\lambda+m-1} & \gamma \geq \theta \\ \frac{\lambda}{\lambda+\frac{m-1}{1+e^\gamma}} & \gamma < \theta \end{cases} \tag{31}$$

Proof. Let $g_i(x)$ be the probability density function for X_i, and let $G_i(x)$ be its cumulative density function. We have

$$g_1(x) = \cdots = g_{m-1}(x) = \begin{cases} \frac{1}{2}e^{-x} + \frac{1}{2}e^{-(x-\theta)} & x \geq \theta. \\ \frac{1}{2}e^{-x} & x < \theta. \end{cases}$$

$$G_1(x) = \cdots = G_{m-1}(x) = \begin{cases} 1 - \frac{1}{2}e^{-x} - \frac{1}{2}e^{-(x-\theta)} & x \geq \theta. \\ \frac{1}{2} - \frac{1}{2}e^{-x} & x < \theta \end{cases}$$

$$g_m(x) = \lambda e^{-\lambda x}$$

$$G_m(x) = 1 - e^{-\lambda x}$$

Thus

$$\begin{aligned}
&\mathbb{P}(X_m = \min\{X_1, \ldots, X_m\} \mid \min\{X_1, \ldots, X_m\} = \gamma) \\
&= \frac{g_m(\gamma)(1 - G_1(\gamma))^{m-1}}{g_m(\gamma)(1 - G_1(\gamma))^{m-1} + (m-1)g_1(\gamma)(1 - G_m(\gamma))(1 - G_1(\gamma))^{m-2}} \\
&= \frac{1}{1 + (m-1)\frac{g_1(\gamma)(1-G_m(\gamma))}{g_m(\gamma)(1-G_1(\gamma))}} \\
&= \begin{cases} \frac{\lambda}{\lambda+m-1} & \gamma \geq \theta \\ \frac{\lambda}{\lambda + \frac{m-1}{1+e^{\gamma}}} & \gamma < \theta \end{cases}
\end{aligned}$$

□

Define

$$\Phi_r := \mathbb{P}(X_m = \min\{X_1, \ldots, X_m\}) = \mathbb{P}(b^* \in B_r^* \mid b^* \notin B_{r-1}^*).$$

If f, F denote the pdf and cdf of $\min\{X_1, \ldots, X_m\}$, then

$$\Phi_r = \int_{\gamma=0}^{\theta} \frac{\lambda f(\gamma)}{\lambda + \frac{m-1}{1+e^{\gamma}}} \mathrm{d}\gamma + \frac{\lambda(1 - F(\theta))}{\lambda + m - 1}.$$

Now assuming that $\gamma \leq \theta$, we have

$$\begin{aligned}
1 - F(\gamma) = \mathbb{P}(\min\{X_1, \ldots, X_m\} \geq \gamma) &= \left(\frac{1 + e^{-\gamma}}{2}\right)^{m-1} e^{-\lambda\gamma} \\
&= \left(\frac{1 + e^{-\gamma}}{2}\right)^{m-1} + O(\lambda).
\end{aligned}$$

Note that above and for the rest of this section, asymptotics are as $\lambda \to 0$ and all other variables are held constant. Taking derivatives, for $\gamma \leq \theta$ we have

$$\begin{aligned}
f(\gamma) &= \frac{1}{2}(m-1)e^{-\gamma}\left(\frac{1 + e^{-\gamma}}{2}\right)^{m-2} e^{-\lambda\gamma} + \left(\frac{1 + e^{-\gamma}}{2}\right)^{m-1} \lambda e^{-\lambda\gamma} \\
&= \frac{1}{2}(m-1)e^{-\gamma}\left(\frac{1 + e^{-\gamma}}{2}\right)^{m-2} + O(\lambda)
\end{aligned}$$

So,

$$\begin{aligned}
\Phi_r &= \int_{\gamma=0}^{\theta} \frac{\lambda \cdot \frac{1}{2}(m-1)e^{-\gamma}\left(\frac{1+e^{-\gamma}}{2}\right)^{m-2}}{\lambda + \frac{m-1}{1+e^{\gamma}}} \mathrm{d}\gamma + \frac{\lambda\left(\frac{1+e^{-\theta}}{2}\right)^{m-1}}{\lambda + m - 1} + O(\lambda^2) \\
&= \frac{\lambda}{2^{m-1}} \int_{\gamma=0}^{\theta} (1 + e^{-\gamma})^{m-1} \mathrm{d}\gamma + \frac{\lambda(1 + e^{-\theta})^{m-1}}{2^{m-1}(m-1)} + O(\lambda^2) \\
&= \frac{\lambda}{2^{m-1}} \int_{x=e^{-\theta}}^{1} \frac{(1+x)^{m-1}}{x} \mathrm{d}x + \frac{\lambda(1 + e^{-\theta})^{m-1}}{2^{m-1}(m-1)} + O(\lambda^2) \\
&= \lambda\Psi(r, \theta) + O(\lambda^2),
\end{aligned}$$

where we have defined (recalling $m = m(r) = n - r + 2$)

$$\Psi(r,\theta) := \frac{1}{2^{m-1}}\left(\int_{x=e^{-\theta}}^{1}\frac{(1+x)^{m-1}}{x}\mathrm{d}x + \frac{(1+e^{-\theta})^{m-1}}{m-1}\right).$$

It follows from (30) that

$$\begin{aligned}
P(n,r) = \lim_{\lambda\to 0^+}\lambda^{-1}\mathbb{P}(b^*\in B_r^*) &= \lim_{\lambda\to 0^+}\lambda^{-1}\left(1-\prod_{i=1}^{r}\mathbb{P}(b^*\notin B_r^* \mid b^*\notin B_{r-1}^*)\right)\\
&= \lim_{\lambda\to 0^+}\lambda^{-1}\left(1-\prod_{i=1}^{r}(1-\Phi_i)\right)\\
&= \lim_{\lambda\to 0^+}\lambda^{-1}\left(1-\prod_{i=1}^{r}\left(1-\lambda\Psi(i,\theta)+O(\lambda^2)\right)\right)\\
&= \sum_{i=1}^{r}\Psi(i,\theta)\\
&= \sum_{i=1}^{r}\frac{1}{2^{n-i+1}}\left(\int_{x=e^{-\theta}}^{1}\frac{(1+x)^{n-i+1}}{x}\mathrm{d}x + \frac{(1+e^{-\theta})^{n-i+1}}{n-i+1}\right) \qquad (32)\\
&=: \Upsilon(r,\theta).
\end{aligned}$$

Now using $L_n(\theta) = C(n,n)$, Lemma 9, and $P(n,r) = \Upsilon(r,\theta)$ we get

$$\mathbb{E}\phi(\theta) = -\theta\alpha n + \mathbb{E}C(n,n) = -\theta\alpha n + \sum_{r=1}^{n}\mathbb{E}(C(n,r)-C(n,r-1)) = -\theta\alpha n + \sum_{r=1}^{n}\frac{\Upsilon(r,\theta)}{r}. \qquad (33)$$

So the relaxation of (23) gives an expected lower bound of $\phi(\theta^*)$ where θ^* is the solution to

$$\begin{aligned}
\alpha n &= \sum_{r=1}^{n}\frac{1}{r}\frac{\partial\Upsilon(r,\theta)}{\partial\theta}\\
&= \sum_{r=1}^{n}\frac{1}{r}\sum_{i=1}^{r}\frac{(1+e^{-\theta})^{n-i}}{2^{n-i+1}}\\
&= \frac{(1+e^{-\theta})^n}{2^n}\sum_{r=1}^{n}\frac{2^r-(1+e^{-\theta})^r}{r(1+e^{-\theta})^r(1-e^{-\theta})}\\
&\sim ne^{-\psi/2}\sum_{r=1}^{n}\frac{e^{\psi r/2n}-1}{r\psi}\\
&\sim n\psi^{-1}e^{-\psi/2}\int_{x=0}^{\psi/2}\frac{e^x-1}{x}\mathrm{d}x
\end{aligned}$$

assuming that $\theta = \psi/n, \psi = O(1)$ (this assumption is justified in Subsection 4.2). Note that while $\psi > 0$ for any $\alpha < 1/2$, the above expression approaches $n/2$ when $\psi \searrow 0$. This verifies our estimate for $\mathbb{E}Z_\alpha$.

4.1 Concentration of ϕ

Concentration of $\phi(\theta)$ around its expectation can be justified as in Section 2.3. All we need is verification that w.h.p. the optimal matchings do not contain edges of weight more than $n^{-9/10}$. This is a fairly standard calculation that depends on the expansion of $K_{n,n,p}$. We will use the following lemma:

Lemma 10. *In $K_{n,n,p}$ with $p = n^{-10/11}$, with probability $1 - o(n^{-200})$, for every matching M of size $n-1$, there is an augmenting path of length at most 53.*

Proof. Indeed suppose that the bipartition of vertices in $K_{n,n}$ is A, B and N_p denotes neighborhood in $K_{n,n,p}$. Then,

$$\begin{aligned}
&\mathbb{P}(\exists S \subseteq A : |S| \le n^{10/11}\log n, |N_p(S)| \le np|S|/2\log n)\\
&\le \sum_{s=1}^{n^{10/11}\log n} \binom{n}{s}\binom{n}{n-nps/2\log n}(1-p)^{(n-nps/2\log n)s}\\
&\le \sum_{s=1}^{n^{10/11}\log n} \left(\frac{ne}{s}\right)^s \left(\frac{ne}{nps/2\log n}\right)^{nps/2\log n} e^{-(n-nps/2\log n)ps} \le \sum_{s=1}^{n^{10/11}\log n} e^{-nps/3} = o(n^{-200}).
\end{aligned}$$

Suppose that $u \in A, v \in B$ are the two vertices not covered by M. The set $A_1 \subseteq A$ of vertices reachable from u by an alternating path of length at most 26 has $|A_1| \ge n/3$, with the required probability. Indeed, starting at u we can find the tree containing all possible such alternating paths by iteratively finding the neighborhood of the current bottom level of the tree, then taking the matching edges covering those vertices, and repeating. Every time we add two levels to the tree, the number of vertices at the bottom level expands by a factor of at least $np/3\log n = n^{1/11}/3\log n$ (where we replaced a "2" with a "3" to account for any vertices that were already in the tree). This lasts until the bottom layer has $n^{10/11}\log n$ vertices, which happens by the time we reach level 26. Two more levels give us another expansion by factor $n^{1/11}/3\log n$, bringing us to $n/3$ vertices in the bottom level. Similarly, there is a set of at least $n/3$ vertices B_1 in B that are reachable from v by an alternating path of length at most 26. Then w.h.p., whatever the sets A_1, B_1, there is a non-matching edge joining A_1 and B_1, giving an alternating path P from u to v of length at most 53. Indeed, the probability that there exist A_1, B_1 not joined by at least $n+1$ edges is at most $2^{2n}e^{-\Omega(n^2p)} = o(n^{-200})$. □

Suppose the minimum cost perfect matching M^* contains an edge $e = (u, v)$, $u \in A, v \in B$, of weight at least $n^{-9/10}$. Remove e to obtain near perfect matching M. Note that $n^{-10/11}$ is the probability that an edge is blue and has cost at most $(2+o(1))n^{-10/11}$. Using the path P promised by Lemma 10, we can find a matching with a cost $n^{-9/10} - O(n^{-10/11})$ less than M^*, a contradiction.

4.2 Bounding θ^*

We need to justify the assumption that $\theta^* = O(1/n)$. The dual function $\phi(\theta) = L_n(\theta) - \theta\alpha n$ and w.h.p. $L_n(\theta) = O(1)$, independent of θ, given that we can just use blue edges for a perfect matching. We have to maximise this over $\theta \ge 0$. This rules out $\theta \ge K/n$ for some constant $K > 0$, as we know that $\max \phi(\theta) \ge \zeta(2)$.

We next estimate $\phi(\theta^*)$ assuming $\theta^* = \psi/n$. For this, we go back to (32) and (33). First of all

$$\begin{aligned}\sum_{r=1}^{n}\sum_{i=1}^{r}\frac{1}{r}\cdot\frac{(1+e^{-\psi/n})^{n-i+1}}{2^{n-i+1}(n-i+1)} &= \sum_{r=1}^{n}\sum_{i=1}^{r}\frac{1}{r}\cdot\frac{(1-\frac{\psi}{2n}+O(n^{-2}))^{n-i+1}}{2^{n-i+1}(n-i+1)}\\ &= \sum_{r=1}^{n}\sum_{i=1}^{r}\frac{e^{-(n-i+1)\psi/2n}+O(1/n)}{r(n-i+1)} \qquad (34)\\ &= \sum_{i=1}^{n}\frac{e^{-i\psi/2n}}{i}\sum_{r=n-i+1}^{n}\frac{1}{r}+o(1)\\ &= -\int_{x=0}^{1}\frac{e^{-\psi x/2}\log(1-x)}{x}\mathrm{d}x+o(1). \qquad (35)\end{aligned}$$

Then we have

$$\begin{aligned}\int_{x=e^{-\psi/n}}^{1}\frac{(1+x)^{n-i+1}}{x}\mathrm{d}x &= \int_{y=0}^{1-e^{-\psi/n}}\frac{\left(1-\frac{y}{2}\right)^{n-i+1}}{1-y}\mathrm{d}y\\ &= \int_{y=0}^{1-e^{-\psi/n}}\frac{e^{-(n-i+1)y/2}}{1-y}\mathrm{d}y+O(1/n)\\ &= \frac{2}{n-i+1}(1-e^{-(n-i+1)\psi/2n})+O(1/n).\end{aligned}$$

Then, using (34), we obtain

$$\sum_{r=1}^{n}\sum_{i=1}^{r}\frac{1}{r}\cdot\frac{2}{n-i+1}(1-e^{-(n-i+1)\psi/2n}) \sim \frac{\pi^2}{3}+2\int_{x=0}^{1}\frac{e^{-\psi x/2}\log(1-x)}{x}\mathrm{d}x+o(1).$$

Combined with (35), we get the expression in (5).

Now $\phi(\theta^*)$ only gives a lower bound on the minimum cost of a perfect matching. We discuss a matching upper bound. Let $\mathcal{M}^*(\theta)$ denote the set of matchings that minimise W_θ. According to [8], $\phi(\theta^*)$ is the objective value of the solution to an LP relaxation of (23) and this solution is a convex combination of matchings in $\mathcal{M}^*(\theta)$. Thus, if $\mathcal{M}^*(\theta^*)$ contains a unique matching M^* then $C(M^*) = \phi(\theta^*)$ and we are done.

Lemma 11. *With probability one, $|\mathcal{M}^*(\theta)| \le 2$ for all $\theta > 0$.*

Proof. Suppose that there are three distinct members of $\mathcal{M}^*(\theta)$ that minimise W_θ. This implies that there are three distinct perfect matchings $M_i, i = 1, 2, 3$ with equal values for $C(M_i) + \theta|M_i \cap E_R|$. Now if $|M_1 \cap E_R| = |M_2 \cap E_R|$ then $C(M_1) = C(M_2)$, an event of probability zero. We can therefore assume that all the $|M_i \cap E_R|$ are distinct. But this implies, that

$$\theta = \frac{C(M_1)-C(M_2)}{|M_2\cap E_R|-|M_1\cap E_R|} = \frac{C(M_1)-C(M_3)}{|M_3\cap E_R|-|M_1\cap E_R|} \qquad (36)$$

or

$$(C(M_1)-C(M_2))(|M_3\cap E_R|-|M_1\cap E_R|) = (C(M_1)-C(M_3))(|M_2\cap E_R|-|M_1\cap E_R|),$$

an event of probability zero. □

So assume there are $M_1, M_2 \in \mathcal{M}^*(\theta^*)$ that satisfy $C(M_1) + \theta^*|M_1 \cap E_R| = C(M_2) + \theta^*|M_2 \cap E_R| = L_n(\theta^*)$. Following [2], we prove

Lemma 12. *Given M_1, M_2 we can construct a matching M with $n-1$ edges for which $C(M) \le \phi(\theta^*)+O(1/n)$ and $|M \cap E_R| \le \alpha n$.*

Proof. Now we cannot have $\delta_1 = \min\{|M_1 \cap E_R|, |M_2 \cap E_R|\} - \alpha n > 0$ else $\phi(\theta^* + \delta_2) > \phi(\theta^*)$ for some small $\delta_2 > 0$. This follows from Lemma 11. Let $\xi = \min\{W_{\theta^*}(M) - W_{\theta^*}(M_1) : M \neq M_1, M_2\} > 0$ and let $\delta_2 = \xi/2n$. Then we have $C_{\theta^*+\delta_2}(M) \ge C_{\theta^*+\delta_2}(M_1) + \xi - n\delta_2 > C_{\theta^*+\delta_2}(M_1)$ for $M \neq M_1, M_2$ so that M_1, M_2 minimise $C_{\theta^*+\delta_2}$. And then $\phi(\theta^* + \delta_2) - \phi(\theta^*) = (|M_1 \cap E_R| - \alpha n)\delta_2 > 0$.

Now $\theta^* > 0$ w.h.p. This is because $\phi(0)$ is the minimum cost of a matching ignoring the color constraint. Then w.h.p. this matching M, which uniquely minimises cost, satisfies $|M \cap E_R| \sim n/2$ and then $\phi(\delta_3) = C(M) + \delta_3(1/2 - o(1) - \alpha)n > C(M) = \phi(0)$ for δ_3 sufficiently small, contradiction.

Given this, we can rule out $\delta_4 = \max\{|M_1 \cap E_R|, |M_2 \cap E_R|\} - \alpha n < 0$. For then we could argue as for $\delta_1 > 0$ that we can reduce θ^* by a small positive amount, less than θ^*, and increase ϕ.

If $|M_i \cap E_R| = \alpha n$ for some i, then M_i itself is feasible and satisfies $C(M_i) = L_n(\theta^*) - \theta^*|M_i \cap E_R| = L_n(\theta^*) - \theta^*\alpha n = \phi(\theta^*)$, so there is nothing to prove.

Assume then that $|M_1 \cap E_R| < \alpha n < |M_2 \cap E_R|$. Let $K = \{e_1, e_2, \dots, e_k\} = M_1 \oplus M_2$, where M_1 is the set of odd subscript edges. Assume that each cycle of $M_1 \oplus M_2$ appears as some contiguous subset $\{e_j, e_{j+1}, \dots, e_{j'}\}$ of K. Let $a_i = \delta(e_i)W_{\theta^*,e_i}$ where $\delta(e_i) = 1$ for $e_i \in M_1$ and -1 otherwise. Then

$$W_{\theta^*}(M_1) - W_{\theta^*}(M_2) = \sum_{i=1}^{k} a_i = 0$$

and so there exists ℓ such that (below, subscripts are taken modulo k)

$$\sum_{j=1}^{t} a_{\ell+j} \ge 0 \text{ for } t = 0, 1, \dots, k-1. \tag{37}$$

Indeed, it suffices to take $\ell = \operatorname{argmin}\left\{\sum_{j=1}^{\ell} a_j : 1 \le \ell \le k\right\}$. This is the content of the gasoline lemma of Lovász [12], Problem 3.21.

For $t \ge 0$ let $X_t = M_1 \cup \{e_{\ell+j} : j = 1, \dots, t,\ \ell + j \text{ is even}\} \setminus \{e_{\ell+j} : j = 1, \dots, t,\ \ell + j \text{ is odd}\}$. Let $\tau = \max\{t : |X_t \cap E_R| \le \alpha n\}$. Then we must have $\ell + \tau$ odd, since $X_{t+1} \subseteq X_t$ when $\ell + t$ is even. Let $M = X_\tau \setminus e_\ell$ if ℓ is odd, and $M = X_\tau \setminus e_{\ell+1}$ if ℓ is even. Note that M is a matching, being obtained from M_1 via an alternating path. Also, $|M \cap E_R| \le |X_\tau \cap E_R| \le \alpha n$ and $|M| = n - 1$. Note that (37) implies that

$$W_{\theta^*}(M) \le W_{\theta^*}(X_\tau) \le W_{\theta^*}(M_1).$$

So,

$$\begin{aligned} C(X_\tau) &= W_{\theta^*}(X_\tau) - \theta^*|X_\tau \cap E_R| \\ &= W_{\theta^*}(X_\tau) - \theta^*\alpha n + \theta^*(\alpha n - |X_\tau \cap E_R|) \\ &\le W_{\theta^*}(M_1) - \theta^*\alpha n + \theta^* \\ &= \phi(\theta^*) + \theta^*. \end{aligned}$$

The maximality of τ implies that $e_{\ell+\tau+1} \in E_R$. So,

$$C(M) \le C(X_\tau) \le \phi(\theta^*) + \theta^* = \phi(\theta^*) + O(1/n). \tag{38}$$

Furthermore, by construction,

$$|M \cap E_R| \leq \alpha n. \tag{39}$$

□

At this point, we need to deal with the fact that $|M| = n - 1$.

Lemma 13. *W.h.p. for every matching M of size $n-1$ there is an augmenting path that creates a matching M' with (i) $|M'| = |M| + 1$, (ii) $C(M') \leq C(M) + O(n^{-10/11})$ and (iii) $|M' \cap E_R| \leq |M \cap E_R|$.*

Proof. Apply Lemma 10 where $K_{n,n,p}$ is defined by the blue edges of cost at most $2n^{-10/11}$. The promised augmenting path will produce a perfect matching. The increase in cost will be $O(n^{-10/11})$ and no red edges will be added. □

Combining Lemmas 12 and 13, w.h.p. there is a perfect matching M' satisfying $|M' \cap E_R| \leq \alpha n$ and $C^* \leq C(M') \leq \phi(\theta^*) + O(n^{-10/11}) + O(1/n) = \phi(\theta^*) + o(1)$. Weak duality gives $C^* \geq \phi(\theta^*)$, and hence $C^* = \phi(\theta^*) + o(1)$ w.h.p. Thus we have that w.h.p.

$$\phi(\theta^*) \leq Z_\alpha \leq \phi(\theta^*) + o(1).$$

So, $Z_\alpha \sim \phi(\theta^*)$ w.h.p. and then the concentration of ϕ around its mean shows that Z_α is also concentrated around the mean defined in the statement of Theorem 3.

5 Proof of Theorem 4

Karp's patching algorithm first solves the problem of finding a minimum weight perfect matching M^* (Assignment Problem) in $K_{n,n}$. We associate this solution to a cycle cover, i.e. a set of vertex disjoint directed cycles in the complete digraph $\vec{K}_n$. (Assume that the vertices of $K_{n,n}$ are $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$. Edge $\{a_i, b_j\}$ of M^* is replaced by the directed edge (i, j) of $\vec{K}_n$. This produces a cycle cover. Note that $\{a_i, b_i\}$ gives rise to a loop at vertex i. Note also that this shows that $Z_\alpha(PM) \leq Z_\alpha(ATSP)$.)

We instead have to solve the problem of finding a minimum weight perfect matching under edge color constraints. We can solve the associated Linear Programming relaxation and then applying Lemma 12 find an asymptotically optimal matching M^* that satisfies the color constraints. It follows from the proof of Lemma 10 that no edge of the matching M^* has $X_e \geq \eta_0 = C_0 n^{-1/3}$ for some $C_0 > 0$. So, w.h.p. we can construct M^* by ignoring all edges E_{Large} of cost greater than η_0.

One can argue by symmetry as was done by Karp [11] that M^* has a uniform distribution and so its associated cycle cover has $m = O(\log n)$ cycles. (Choose the constrained optimum matching using a label-equivariant tie-breaking rule, or choose uniformly among all constrained optima. Since the joint law of colors and costs is invariant under permutations of the two vertex classes, the resulting perfect matching is uniformly distributed over all perfect matchings.) Let these cycles be $K_1, K_2, \ldots, K_m$ where $|K_i| \geq |K_{i+1}|$. Let $E_1, E_2, \ldots, E_{m-1}$ be a random partition of E_{Large} into $m-1$ almost equal size subsets. We proceed by patching the cycles together. Given cycles K, K' we patch them together by finding edges $e = (a, b) \in K, f = (c, d) \in K'$ and then replacing e, f by $e' = (a, d), f' = (c, b)$. This gives us a new cycle with vertex set $V(K) \cup V(K')$. We call e, f a patching pair if $(a, d), (c, b)$ are blue and both edges have cost in $[\eta_0, 2\eta_0]$. (If one cycle is a loop on vertex a then we modify the construction appropriately, replacing an edge (b, c) in the larger cycle by the

path b, a, c.) Different pairs of arcs in K_i, K_j give rise to disjoint patching pairs with independent edge costs. We ignore the saving associated with deleting e, f and only look at the extra cost $C(e') + C(f')$ incurred.

We use E_1 to find a cycle to patch into K_1, creating a cycle $\widehat{K}$. In general, for $i \geq 2$, we use E_i to find a cycle to patch into $\widehat{K}$, increasing its size. Assuming that we search over all possible patching pairs involving $\widehat{K}$ and unabsorbed cycles, the number of possible patching pairs $|\widehat{K}|(n - |\widehat{K}|) \geq n - 1$ holds for every i. So the probability we fail to find the $m - 1$ needed patches is at most $m(1 - \Omega(\eta_0^2/m^2))^{n-1} = o(1)$. The additional cost associated with this construction is $O(mn^{-1/3}) = o(1)$ and only blue edges are added and so the red count does not increase. Thus the resulting Hamilton cycle is feasible and has cost $Z_\alpha(PM) + o(1)$. Since $Z_\alpha(PM) \leq Z_\alpha(ATSP)$, this proves Theorem 4.

6 General distributions

We will now consider the situation where the costs/weights X_e are independent copies of a continuous random variable C. We assume that C has a continuous density f and satisfies

(i) $f(x) = a + bx + O(x^2)$ for $0 \leq x \leq L$, where a, b are constants and $aL \geq 1$.

(ii) $\mathbb{P}(C \geq x) \leq c_1 e^{-c_2 x}$ for constants $c_1, c_2 > 0$.

(iii) To avoid some pathologies, we will also assume that there is a constant M such that $f(x) \leq aM$ for $x \geq 0$.

The prime examples are the uniform $[0, 1]$ distribution $U[0, 1]$ ($a = 1, b = 0, M = 1$) and the exponential mean 1 distribution $EXP(1)$ ($a = 1, b = -1, c_1 = c_2 = M = 1$).

After replacing C by aC, the new density at zero is 1. Thus it suffices to prove the result for $a = 1$; the original costs are obtained by dividing the limiting constants by a.

First consider Theorem 1 where we assumed uniform costs. Let $F(x) = \mathbb{P}(C \leq x) = x + \frac{b}{2}x^2 + O(x^3)$ and let $\widehat{C}$ be uniformly distributed. Then we can take $C = F^{-1}(\widehat{C})$ and C will have the correct distribution. If $\widehat{C} = y$ then $C = x = y + O(y^2)$. Now Lemma 2 implies that w.h.p. the maximum edge cost in our solution is $O(n^{-9/10})$, we see that the optimal costs using $C, \widehat{C}$ differ by $O(n \times n^{-9/5}) = o(1)$. Thus Theorem 1 holds with more generality.

Now consider Theorems 2, 3 where we use $EXP(1)$ for costs/weights. This time we take $\widehat{C} = -\log(1 - F(C))$ so that $\mathbb{P}(\widehat{C} \leq x) = \mathbb{P}(-\log(1 - F(C)) \leq x) = \mathbb{P}(F(C) \leq 1 - e^{-x}) = 1 - e^{-x}$. Notice that $-\log(1 - F(x)) = x + O(x^2)$. In Theorem 2 we only use edges of cost $O(\log n/n)$ and in Theorem 3 we only use edges of weight $O(n^{-9/10})$ and so we can proceed as for Theorem 1.

In the case of Theorem 4 we use the same $\widehat{C}$ as for Theorems 2 and 3. In which case the possible increases in error are $O(n \cdot n^{-9/5})$ for the Assignment Problem and $O(\log n \cdot n^{-2/3})$ for the patching of the cycles.

References

[1] D. J. Aldous, The $\zeta(2)$ limit in the random assignment problem, *Random Structures Algorithms* 18 (2001), no. 4, 381–418

[2] A. Berger, V. Bonifaci, F. Grandoni and G. Schaefer, Budgeted matching and budgeted matroid intersection via the gasoline puzzle, *Lecture Notes in Computer Science* 5035 (2007) 273-287.

[3] A. Beveridge, A.M. Frieze and C. McDiarmid, Random minimum length spanning trees in regular graphs, *Combinatorica* 18 (1998) 311-333.

[4] E. Dynkin, Theory of Markov processes, Pergamon Press, 1961.

[5] J. Edmonds, Submodular functions, matroids and certain polyhedra, in *Combinatorial Structures and their Applications, R. Guy et al, eds., Gordon and Breach* (1970) 69-87.

[6] A. M. Frieze, On the value of a random minimum spanning tree problem, Discrete Appl. Math. **10** (1985), no. 1, 47–56

[7] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[8] A.M. Geoffrion, Lagrangian relaxation for integer programming. *Mathematical Programming Study* 2 (1974) 82-114.

[9] S. Janson, The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph, *Random Structures Algorithms* 7 (1995) 337-355.

[10] S. Janson, One, two and three times $\log n/n$ for paths in a complete graph with random weights, Combin. Probab. Comput. **8** (1999), no. 4, 347–361

[11] R.M. Karp, A patching algorithm for the non-symmetric traveling salesman problem, *SIAM Journal on Computing* 8 (1979) 561-573.

[12] L. Lovász, Combinatorial Problems and Exercises, AMS Chelsea Publishing, 2nd Edition, 2007.

[13] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* 81 (1995) 73–205.