On random multi-dimensional assignment problems

Alan Frieze* Wesley Pegden† Tomasz Tkocz
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA15217
U.S.A.

November 25, 2020

Abstract

We study random multidimensional assignment problems where the costs decompose into the sum of independent random variables. In particular, in three dimensions, we assume that the costs $W_{i,j,k}$ satisfy $W_{i,j,k} = a_{i,j} + b_{i,k} + c_{j,k}$ where the $a_{i,j}, b_{i,k}, c_{j,k}$ are independent uniform $[0, 1]$ random variables. Our objective is to minimize the total cost and we show that w.h.p. a simple greedy algorithm is a $(3 + o(1))$-approximation. This is in contrast to the case where the $W_{i,j,k}$ are independent exponential rate 1 random variables. Here all that is known is an $n^{o(1)}$-approximation, due to Frieze and Sorkin.

1 Introduction

The (planar) three dimensional assignment problem is a natural generalisation of the classical assignment problem. As an optimization problem it can be expressed as follows: we are given real values $W_{i,j,k}$ for $i, j, k \in [n]$ and we are asked to

$$Minimize \left\{ \sum_{i=1}^{n} W_{i,\sigma(i),\tau(i)} : \ \sigma, \tau \text{ are permutations of } [n] \right\}.$$

*Research supported in part by NSF grant DMS1661063
†Research supported in part by NSF grant DMS1363136

This is an NP-hard problem and occurs for example as a practical problem [3]. In this paper we study the following simple greedy heuristic:

Algorithm 1 GREEDY(m)

1: Let $B := C := [n]$, and $T := \emptyset$.
2: **for** $i = 1, \ldots, m$ **do**
3: Let $W_{i,j,k} = \min \{W_{i,j',k'} : j' \in B, k' \in C\}$.
4: Add (i, j, k) to T and remove j from B and k from C.
5: Return the set of triples in T as a partial assignment.
6: Complete the assignment with one of the remaining $(n-m)!^2$ possibilities.

Several authors have considered the average case where the $W_{i,j,k}$ are random variables. Kravtsov [4] considered the case where the $W_{i,j,k}$ are chosen randomly from $[1, M]$ where $M = n^\alpha$ for some $\alpha < 1$. Here the minimum is at least n and it is not difficult to show (see Section 4 that with the choice of $m = n - \log n$ that w.h.p. (i) GREEDY(m) runs in polynomial time and (ii) it outputs a solution of value $n + o(n)$. In this case Step 6 can be completed via the choice of an arbitrary completion.

It is more difficult to analyse the case where $M \gg n$ and the case where the $W_{i,j,k}$ are independent exponential rate 1 random variables is (essentially) a scaled version of such a case. This case was considered by Frieze and Sorkin [2] and they proved the following theorem.

Theorem 1 (Frieze and Sorkin)**.** *Suppose that the $W_{i,j,k}$ are independent EXP(1) random variables and that Z_n denote the value of the optimum. Then (a) $\frac{1}{n} \leq \mathbf{E}(Z_n) = O\left(\frac{\log n}{n}\right)$ and (b) there is a polynomial time algorithm that w.h.p. finds a solution of value $\frac{1}{n^{1-o(1)}}$.*

A recent result of Frankston, Kahn, Narayana and Park [1] shows that $\mathbf{E}(Z_n) = O\left(\frac{1}{n}\right)$. This is where the problem stands for such $W_{i,j,k}$ and here we consider the case where

$$W_{i,j,k} = a_{i,j} + b_{i,k} + c_{j,k},\ 1 \leq i, j, k \leq n, \tag{1}$$

where the $a_{i,j}, b_{i,k}, c_{j,k}$ are independent uniform $[0, 1]$ random variables.

We note that the problem considered in [3] was of the form given in (1). We will prove the following theorem.

Theorem 2. *There exists an explicitly defined constant c_1 such that (a) $\mathbf{E}(Z_n) \geq c_1 n^{1/3}$ and (b) GREEDY($n - n^{1/4}$) finds a solution of expected value at most $(3 + o(1))c_1 n^{1/3}$. In this case Step 6 can be completed by choosing an arbitrary completion.*

Before giving a proper proof, we give a heuristic argument for (a). Fix i and consider $W_{i,j,k}$. For $W_{i,j,k}$ to be of order $n^{-\alpha}$ say we need each of 3 uniform $[0, 1]$ variables $a_{i,j}, b_{i,k}, c_{j,k}$ to be of order $n^{-\alpha}$. This happens with probability $O(n^{-3\alpha})$ and there are n^2 choices and $3\alpha = 2$ gives the largest value for α. Summing over i gives (a).

We discuss the rigorous proof of Theorem 2 in Section 2 and in Section 3 we consider the extension to higher dimensions.

1.1 Preliminaries

We sometimes refer to the Hoeffding bounds for the $S = S_1+S_2+\ldots+S_N$ where $S_1, S_2, \ldots, S_N \in [0,1]$ are independent and $\mathbf{E}(S_1) + \mathbf{E}(S_2) + \cdots + \mathbf{E}(S_N) = N\mu$:

$$\mathbf{Pr}(|S - N\mu| \geq \varepsilon N\mu) \leq 2e^{-\varepsilon^2 N\mu/3}. \qquad (2)$$

We say that a sequence of events $\mathcal{E}_n$ occur *quite surely* if $\mathbf{Pr}(\neg\mathcal{E}_n) = O(n^{-K})$ for any constant $K > 0$.

2 Proof of Theorem 2

We begin by analysing the distribution of the smallest weight element in $\Pi_i = \{i\} \times [n]^2$.

2.1 Weights in a fixed plane

In the following we are considering the set of triples (i, j, k) where i is fixed. Then $a_j, b_k, c_{j,k}$ stand for $a_{i,j}, b_{i,k}$ and $c_{j,k}$. Let

$$W_n = \min \left\{a_j + b_k + c_{j,k} : a_j, b_k, c_{j,k},\ j, k \in [n] \text{ are independent uniform } [0,1] \text{ random variables}\right\}.$$

Lemma 3. $\mathbf{E}(W_n) \approx c_1 n^{-2/3}$, *where* $c_1 = 6^{1/3}\Gamma(3/4)$, *where* Γ *denotes Euler's Gamma function.*

Proof. Let

$$L = \log n,\ J = \left\{j : a_j \leq \frac{L}{n^{2/3}}\right\},\ K = \left\{k : b_k \leq \frac{L}{n^{2/3}}\right\},$$
$$X = \left\{(j,k) \in J \times K : c_{j,k} \leq \frac{L}{n^{2/3}}\right\}. \qquad (3)$$

Now $|J|, |K|$ are both distributed as the Binomial $B(n, L/n^{2/3})$ and so it follows from (2) that

$$|J|, |K| \in \left[\frac{1}{2}Ln^{1/3}, \frac{3}{2}Ln^{1/3}\right] \quad q.s. \tag{4}$$

Conditional on the sizes of J, K we have $|X|$ is distributed as $B(|J| \cdot |K|, L/n^{2/3})$. It follows from (2) with $\varepsilon = 6/7$ and (4) that

$$\begin{aligned}\mathbf{Pr}\left(|X| \notin \left[\frac{L^3}{8}, 10L^3\right]\right) &\leq \mathbf{Pr}(\neg(4)) + \mathbf{Pr}\left(B\left(\frac{L^2n^{2/3}}{4}, \frac{L}{n^{2/3}}\right) \leq \frac{L^3}{8}\right) \\ &+ \mathbf{Pr}\left(B\left(\frac{9L^2n^{2/3}}{4}, \frac{L}{n^{2/3}}\right) \geq 10L^3\right) \leq e^{-\Omega(L^3)}. \quad (5)\end{aligned}$$

Thus let $\mathcal{E}_L$ denote the event that $|X| \in \left[\frac{L^3}{8}, 10L^3\right]$, which from (5) occurs q.s.

Let $\mathcal{E}_M$ denote the event that the edges in X **almost** form a matching. By this we mean that the graph induced by X consists of a matching M plus at most 4 extra edges Y. Then, because

$$\mathbf{E}(W_n) = \mathbf{E}(W_n \mid \mathcal{E}_M)\mathbf{Pr}(\mathcal{E}_M) + \mathbf{E}(W_n \mid \neg\mathcal{E}_M)\mathbf{Pr}(\neg\mathcal{E}_M),$$

we see that

$$\mathbf{E}(W_n \mid \mathcal{E}_M)\mathbf{Pr}(\mathcal{E}_M) \leq \mathbf{E}(W_n) \leq \mathbf{E}(W_n \mid \mathcal{E}_M) + 3\mathbf{Pr}(\neg\mathcal{E}_M), \tag{6}$$

where we use the fact that $W_n \leq 3$, always.

We first deal with $\mathbf{Pr}(\neg\mathcal{E}_M)$ by showing that.

$$\mathbf{Pr}(\neg\mathcal{E}_M) = O\left(\frac{L^{15}}{n}\right). \tag{7}$$

Let

$$p = \frac{L}{n^{2/3}}.$$

Condition on J, K satisfying (4). Let Γ_X be the bipartite graph induced by X and note that it is distributed as the binomial random graph $G_{|J|,|K|,p}$.

Claim 1. *The following holds with probability $1 - O(L^{15}/n)$: (i) Γ_X has no connected component with 4 or more edges and (ii) Γ_X has at most one component with 3 edges and (iii) Γ_X has at most 2 components with 2 edges.*

Proof of claim: Let $K = |J| + |K|$.

$$\mathbf{Pr}(\neg(i)) = O\left(\left(\binom{K}{4} + \binom{K}{5}\right)p^4\right) = O\left(\frac{L^9n^{5/3}}{n^{8/3}}\right) = O\left(\frac{L^9}{n}\right).$$

$$\mathbf{Pr}(\neg(ii)) = O\left(\left(\binom{K}{4}p^3\right)^2\right) = O\left(\frac{L^{14}n^{8/3}}{n^4}\right) = O\left(\frac{L^{14}}{n^{4/3}}\right).$$

$$\mathbf{Pr}(\neg(iii)) = O\left(\left(\binom{K}{3}p^2\right)^3\right) = O\left(\frac{L^{15}n^3}{n^4}\right) = O\left(\frac{L^{15}}{n}\right).$$

End of proof of claim.

Now given $\mathcal{E}_M$ we let $\widehat{W}_n$ denote the minimum weight in M where the weight of edge $(j,k) \in X$ is given by $a_j + b_k + c_{j,k}$. We see that $\widehat{W}_n$ is the minimum of $|M|$ independent copies of $U = (U_1 + U_2 + U_3)p$ where U_1, U_2, U_3 are independent uniform $[0,1]$.

Thus

$$\phi(u) = \mathbf{Pr}(U \geq pu) = 1 - \frac{1}{6}\sum_{k=0}^{\lfloor u \rfloor}(-1)^k \binom{3}{k}(u-k)^3.$$

It follows that

$$\begin{aligned}
\mathbf{E}(\widehat{W}_n \mid \mathcal{E}_L, \mathcal{E}_M, |M|) &= p\int_{u=0}^{3} \mathbf{Pr}(\widehat{W}_n \geq up \mid \mathcal{E}_L, \mathcal{E}_M, |M|)du \\
&= p\int_{u=0}^{3} \phi(u)^{|M|}du \\
&= p(I_1 + I_2 + I_3), \qquad (8)
\end{aligned}$$

where, with $\alpha = 2/3$, we have

$$\begin{aligned}
I_1 &= \int_{u=0}^{1}\left(1 - \frac{u^3}{6}\right)^{|M|} du \qquad (9)\\
&= \int_{u=0}^{L^{-\alpha}}\left(1 - \frac{u^3}{6}\right)^{|M|} du + \int_{u=L^{-\alpha}}^{1}\left(1 - \frac{u^3}{6}\right)^{|M|} du \qquad (10)\\
&= \int_{u=0}^{L^{-\alpha}} \exp\left\{-|M|u^3/6 + O(|M|u^6)\right\} du + O(e^{-\Omega(|M|L^{-3\alpha})})\\
&= \left(1 + O\left(\frac{1}{L^{3\alpha}}\right)\right)\int_{u=0}^{L^{-\alpha}} e^{-|M|u^3/6}du + O(e^{-\Omega(|M|L^{-3\alpha})})\\
&= \left(1 + O\left(\frac{1}{L^{3\alpha}}\right)\right)\int_{u=0}^{\infty} e^{-|M|u^3/6}du\\
&= \frac{2\left(1 + O(L^{-3\alpha})\right)}{6^{2/3}|M|^{1/3}}\int_{x=0}^{\infty} x^{-2/3}e^{-x}dx\\
&= \frac{\left(6^{1/3}\Gamma(4/3) + O(L^{-3\alpha})\right)}{|M|^{1/3}}.
\end{aligned}$$

To rewrite (10) we have used $1-x = e^{-x+O(x^2)}$ to handle the first integral and have bounded the second one by $(1-L^{-\alpha})(1-L^{-3\alpha}/6)^{|M|} \leq e^{-L^{-3\alpha}|M|/6}$. Now because $\phi(u)$ dereases monotonically with u we have

$$I_2 = \int_{u=1}^{2} \phi(u)^{|M|} du \leq \left(\frac{5}{6}\right)^{|M|} \text{ and } I_3 = \int_{u=2}^{3} \phi(u)^{|M|} du \leq \left(\frac{5}{6}\right)^{|M|}.$$

Thus,

$$\mathbf{E}(\widehat{W}_n \mid \mathcal{E}_L, \mathcal{E}_M, |M|) = \frac{\left(6^{1/3}\Gamma(4/3) + O(L^{-2})\right)}{|M|^{1/3}} p. \tag{11}$$

Integrating $|M|$ from (11) we obtain

$$\mathbf{E}(\widehat{W}_n \mid \mathcal{E}_L, \mathcal{E}_M) = \left(6^{1/3}\Gamma(4/3) + O(L^{-2})\right) \times \mathbf{E}((Bin(|J|\cdot|K|, p) - O(1))^{-1/3}) \times p. \tag{12}$$

Given $\mathcal{E}_L$ we see that the binomial is q.s. much greater than 4. Now, for binomial parameters N, q such that Nq large, we have, from (2), that for $\varepsilon > 0$,

$$\begin{aligned}
\mathbf{E}((Bin(N,q) - O(1))^{-1/3}) &= \sum_{k=5}^{N} \binom{N}{k} q^k (1-q)^{N-k} (k - O(1))^{-1/3} \\
&= \sum_{k=(1-\varepsilon)Nq}^{(1+\varepsilon)Nq} \binom{N}{k} q^k (1-q)^{N-k} (k - O(1))^{-1/3} + 2e^{-\varepsilon^2 Nq/3} \\
&= \left(1 + O\left(\frac{1}{Nq}\right)\right) \frac{1}{(Nq)^{1/3}} \sum_{k=(1-\varepsilon)Nq}^{(1+\varepsilon)Nq} \binom{N}{k} q^k (1-q)^{N-k} + 2e^{-\varepsilon^2 Nq/3} \\
&= \frac{1 + O((Nq)^{-1})}{(Nq)^{1/3}} + O(e^{-\varepsilon^2 Nq/3}),
\end{aligned} \tag{13}$$

It then follows from (12) that

$$E(\widehat{W}_n \mid \mathcal{E}_L, \mathcal{E}_M) \approx \frac{6^{1/3}\Gamma(4/3)p}{(|J|\cdot|K|\cdot p)^{1/3}} \tag{14}$$

Arguing as for (13) and using the independence and concentration of $|J|, |K|$ around $Ln^{1/3}$, we see that

$$E(\widehat{W}_n \mid \mathcal{E}_M) \approx 6^{1/3}\Gamma(4/3)p^{2/3}\mathbf{E}\left(\frac{1}{(|I|\cdot|J|)^{1/3}} |\mathcal{E}_M\right) \approx 6^{1/3}\Gamma(4/3)p^{2/3}\frac{1}{(Ln^{1/3})^2} = \frac{6^{1/3}\Gamma(4/3)}{n^{2/3}}. \tag{15}$$

We now have to deal with the at most 4 edges in Y, since $W_n = \min\left\{\widehat{W}_n, Z\right\}$ where Z is the minimum of at most 4 copies of $(U_1 + U_2 + U_3)p$, where U_1, U_2, U_3 are i.i.d. $U[0,1]$. Clearly $\mathbf{E}(W_n) \leq \mathbf{E}(\widehat{W}_n)$ and we need to argue that it is not much smaller. So, let $\mathcal{A} =$

$\left\{\widehat{W}_n \le pL^{-1/2} \le Z\right\}$. Now we have $\mathbf{Pr}(\mathcal{A}) = 1 - O(L^{-1/2})$ and $\mathbf{E}(W_n) \ge \mathbf{E}(\widehat{W}_n \mid \mathcal{A})\mathbf{Pr}(\mathcal{A})$ and so we only have to verify now that $\mathbf{E}(\widehat{W}_n \mid \mathcal{A})$ is asymptotically equal to $\mathbf{E}(\widehat{W}_n)$. Let $\mathcal{L} = \left\{\widehat{W}_n \le pL^{-1/2}\right\}$ and $\mathcal{B} = \left\{pL^{-1/2} \le Z\right\}$. Now because $\widehat{W}_n$ and Z are independent, we have, given $|M|$,

$$\mathbf{E}(\widehat{W}_n \mid \mathcal{A}) = \frac{1}{\mathbf{Pr}(\mathcal{A})}\int_{u=0}^{pL^{-1/2}} \mathbf{Pr}\left(\left\{u \le \widehat{W}_n \le pL^{-1/2}\right\} \cap \mathcal{B}\right) du =$$

$$\frac{1}{\mathbf{Pr}(\mathcal{L})\mathbf{Pr}(\mathcal{B})}\int_{u=0}^{pL^{-1/2}} \left(\mathbf{Pr}\left\{\widehat{W}_n \ge u\right\} - \mathbf{Pr}(\widehat{W}_n \ge pL^{-1/2})\right)\mathbf{Pr}(\mathcal{B})du =$$

$$= \frac{1}{\mathbf{Pr}(\widehat{W}_n \le pL^{-1/2})}\int_{u=0}^{pL^{-1/2}} \mathbf{Pr}(\widehat{W}_n \ge u)du - pL^{-1/2}\frac{\mathbf{Pr}(\widehat{W}_n > pL^{-1/2})}{\mathbf{Pr}(\widehat{W}_n \le pL^{-1/2})}. \quad (16)$$

Now

$$\mathbf{Pr}(\widehat{W}_n > pL^{-1/2}) = \left(1 - \frac{(L^{-1/2})^3}{6}\right)^{|M|} \le e^{-|M|L^{-1/6}/6}.$$

Furthermore,

$$\mathbf{Pr}(\widehat{W}_n \ge u) \ge \left(1 - \frac{u^3}{6}\right)^{|M|}$$

and so integral in the first term of (16) is at least

$$p\int_{p=0}^{L^{-1/2}} \left(1 - \frac{u^3}{6}\right)^{|M|} du.$$

Thus

$$\mathbf{E}(\widehat{W}_n \mid \mathcal{A}) \ge (1 - o(1))p\int_{p=0}^{L^{-1/2}} \left(1 - \frac{u^3}{6}\right)^{|M|} du - e^{-|M|L^{-1/6}/6}$$

and we can proceed as for our estimation of the first integral in (10), this time taking $\alpha = 1/2$.

The lemma now follows after applying (6) and (7). □

This proves Part (a) of Theorem 2, since clearly, $\mathbf{E}(Z_n) \ge n\mathbf{E}(W_n)$.

2.2 Analysis of Greedy

Let now W_m denote the the weight of the triple (i, j, k) added in the mth round of greedy.

Lemma 4. *If $m \le n - n^{1/4}$ then*

$$\mathbf{E}(W_m) \lesssim c_1(n - m + 1)^{-2/3}. \quad (17)$$

Proof. We let J_m, K_m be as defined in (3), where we replace n in the definition by $\nu_m = n - m + 1$. We keep L as $\log n$ though and replace p by $p_m = \frac{L}{\nu_m}$. The values $a_{m,j}, b_{m,k}$ are independent of the first $m-1$ rounds of GREEDY. Now $|J|_m, |K_m|$ are distributed as $Bin(\nu_m, L\nu_m^{-2/3})$ and equation (2) implies that (4) holds q.s. with n replaced by ν_m. Next define X_m iteratively via $X_0 = \emptyset$ and

$$X_m = \left\{ (j,k) \in (J_m \times K_m) \setminus \bigcup_{l<m} X_l,\ c_{j,k} \leq \frac{L}{\nu_m^{2/3}} \right\}.$$

We will show below that

$$\mathbf{Pr}\left(\left| (J_m \times K_m) \cap \bigcup_{l<m} X_l \right| \geq 400L\nu_m^{1/3} \right) = o(n^{-3}). \tag{18}$$

Observe that $c_{j,k}$ for $(j,k) \in X_m$ is unconditioned by the history of GREEDY to this point. Indeed, we will not have needed to expose its value in order to compute the sequence $W_1, W_2, \ldots, W_{m-1}$. But if (18) holds then the analysis of Section 2.1 implies that

$$\mathbf{E}(W_m) \approx c_1 \nu_m^{-2/3}.$$

Indeed, going back to (12) we

$$\text{replace } \mathbf{E}(Bin(|J| \cdot |K|, p)^{-1/3}) \text{ by } \mathbf{E}((Bin(|J_m| \cdot |K_m| - 400L\nu_m^{1/3}, p_m))^{-1/3})$$

and continue as before.

It remains to verify (18). Thus let $Y_m = (J_m \times K_m) \cap \bigcup_{l<m} X_l$ and $Z = |Y_m|$. Now the sequence of choices $J_\ell, K_\ell, \ell \leq m$ are independent and then for $(x,y) \in J_m \times K_m$ and $\ell < m$ we have

$$\mathbf{Pr}((x,y) \in J_\ell \times K_\ell \mid (\ell, x, y) \text{ not added to } T \text{ in Step 4}) \leq \frac{\mathbf{Pr}((x,y) \in J_\ell \times K_\ell)}{\mathbf{Pr}((\ell,x,y) \text{ not added to } T \text{ in Step 4})} \leq \frac{\nu_\ell^{-4/3}}{1 - o(\nu_\ell^{-2})}. \tag{19}$$

It follows (using (4)) that

$$\mathbf{E}(Z) \leq 4L^2\nu_m^{2/3} \sum_{\ell=1}^{m-1} \frac{1}{\nu_\ell^{4/3}} \leq 13L^2\nu_m^{1/3}. \tag{20}$$

Unfortunately, this is not good enough to prove (18). Instead, suppose that $S = \{(x_i, y_i), i \in [s]\} \subseteq J_m \times K_m$ where $s = O(1)$ and S is a matching. Then,

$$\mathbf{Pr}(S \subseteq Y_m) \leq \sum_{i_1 \leq \cdots \leq i_s} \mathbf{Pr}\left(\bigcap_{t=1}^{s} \{(x_t, y_t) \in X_{i_t}\} \right) =$$

$$\sum_{i_1 \leq \cdots \leq i_s} \prod_{t=1}^{s} \mathbf{Pr}\left((x_t, y_t) \in X_{i_t} \middle| \bigcap_{\tau=1}^{t-1} \{(x_\tau, y_\tau) \in X_{i_\tau}\} \right) \leq \sum_{i_1 \leq \cdots \leq i_s} \prod_{t=1}^{s} ((1+o(1))\nu_{i_t}^{-4/3})$$
$$\leq \prod_{t=1}^{s} \sum_{l=1}^{m} \frac{1+o(1)}{(n-l+1)^{4/3}} \leq \left(\frac{4}{(n-m)^{1/3}} \right)^s.$$

Thus,

$$\mathbf{Pr}(\exists \text{ matching } S, |S| = s \mid (4)) \leq \binom{10L^3}{s} \left(\frac{4}{n^{1/12}} \right)^s \leq \left(\frac{40eL^3}{sn^{1/2}} \right)^s = o(n^{-3}) \qquad (21)$$

if $s = 8$. Finally observe that if the maximum size of $S = s \leq 8$ and $|J_m|, |K_m| \leq 10L\nu_m^{1/3}$ then $|Y_m| \leq s(|J_m| + |K_m|) \leq 10sL\nu_m^{1/3}$ and the condition in (18) holds. □

Given Lemma 4 we see that the expected cost of the assignment produced by GREEDY is at most

$$(c_1 + o(1)) \sum_{m=1}^{n-n^{1/4}} \frac{1}{(n-m+1)^{2/3}} + n^{1/4} \approx 3c_1 n^{1/3}. \qquad (22)$$

The final $n - n^{1/4}$ steps cost at most 3 per step and this completes the proof of Theorem 2.

3 Higher Dimensions

Consider for example 4 dimensions. Here we have two reasonable options.

1. $W_{i,j,k,l} = a_{i,j} + b_{i,k} + c_{i,l} + d_{j,k} + e_{j,l} + f_{k,l}$.
2. $W_{i,j,k,l} = a_{i,j,k} + b_{i,j,l} + c_{i,k,l} + d_{j,k,l}$.

We have not considered the first option. The second option is a strightforward generalisation of what we have done so far. Here we will sketch a proof as a series of bullet points that the optimum and the greedy solution for the d-dimensional problem grow at rate $n^{1/d}$ in expectation. By the d-dimensional problem we mean

$$Minimize \left\{ \sum_{i=1}^{n} W_{i,\sigma_1(i),\ldots,\sigma_{d-1}(i)} : \ \sigma_1, \ldots, , \sigma_{d-1} \text{ are permutations of } [n] \right\}. \qquad (23)$$

where

$$W_{i_1,\ldots,i_d} = \sum_{j=1}^{d} A^{(j)}_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_d} \text{ is the sum of independent uniform } [0,1] \text{ random variables.}$$

Let Z_n denote the optimal objective function value in problem (23). We claim that Theorem 2 can be generalised to

Theorem 5. *Suppose that $d \geq 3$. Then there exist constants c_d, C_d such that (a) $\mathbf{E}(Z_n) \geq c_d n^{1/d}$ and (b)* GREEDY*$(n^{1/(d+1)})$ finds a solution of expected value at most $d(c_d + o(1))n^{1/d}$. In this case Step 6 can be completed by choosing an arbitrary completion.*

Proof Sketch:
We can follow the argument in Lemma 3 essentially replacing $n^{1/3}$ by $n^{1/d}$ and $n^{2/3}$ by $n^{(d-1)/d}$. In effect, we make the following replacements:

(a): p becomes $L/n^{(d-1)/d}$.

(b): J, K will be replaced by $I_2, \ldots, I_d$ of expected size np. Here I_j is a representative of a set $\left\{(i_2, i_3, \ldots, i_{j-1}, i_{j+1}, \ldots, i_d, i : A^{(j)}_{i,i_2,\ldots,i_{j-1},i_{j+1},\ldots,i_d}) \leq p\right\}$, for some fixed plane $i_1 = i$.

(c): In which case X becomes $\{\mathbf{j} \in (I_2 \times \cdots \times I_d) : W_{i,\mathbf{j}} \leq p\}$.

(d): (5) becomes $|X| = \Theta(L^d)$. Note that each I_j has distribution $B(n^{d-1}, p)$ and X has distribution $B(|I_2| \times |I_3| \times \cdots \times |I_d|, p)$.

(e): A matching now means a matching in a random $(d-1)$-uniform hypergraph H induced by $I_2 \times \cdots \times I_d$, where possible edges are included with probability p. In the proof of Claim 1, we now let $K = |I_2| + \cdots + |I_d|$, which q.s. is of order $Ln^{1/d}$. We now claim that with probability $\frac{1}{n^{1-o(1)}}$ there are at most $\frac{d}{\ell - 1}$ components of H with $\ell \leq d+1$ edges and no components with $d+2$ or more edges. Indeed, the probability that there are a components of H with ℓ edges can be bounded by

$$\left(\binom{K}{\ell(d-2)+1} p^{\ell}\right)^a = O\left(\frac{L^{\ell(d-1)+1}}{n^{(\ell-1)/d}}\right)^a. \tag{24}$$

Explanation: A minimal size component with ℓ edges is a *cactus* which has $\ell(d-1)+1$ vertices. Thus the LHS of (24) bounds the probability of the existence of a components with at least ℓ edges.

This verifies the claim and shows that if $\mathcal{E}_M$ is the event that X defines a matching M plus $O(1)$ edges, then $\neg\mathcal{E}_M$ is unlikely enough so that we can use (6).

(f): We let $\widehat{W}_n$ denote the minimum weight in M. We see that $\widehat{W}_n$ is the minimum of $|M|$ independent copies of $U = (U_1 + U_2 + \cdots + U_d)p$ where $U_1, U_2, \ldots, U_d$ are independent uniform $[0, 1]$. The sum $p(I_1 + I_2 + I_3)$ becomes $p(I_1 + \cdots + I_d)$ where $I_t = \int_{u=t-1}^{t} \phi(u)^{|M|} du$ and

$$\phi(u) = \mathbf{Pr}(U \geq pu) = 1 - \frac{1}{d!} \sum_{k=0}^{\lfloor u \rfloor} (-1)^k \binom{d}{k} (u-k)^d.$$

The $pI_j, j \geq 2$ are dominated by pI_1 where

$$I_1 = \int_{u=0}^{1} \left(1 - \frac{u^d}{d!}\right)^{|M|} du \approx \frac{1}{|M|^{1/d}} \int_{x=0}^{\infty} e^{-x^d/d!} dx = \frac{(d!)^{1/d}\Gamma(1+1/d)}{|M|^{1/d}}.$$

Domination comes from the fact that $\phi(u) \leq 1 - \frac{1}{d!}$ for $u \geq 1$.

(g): After this we find that (14) becomes

$$E(W_n \mid \mathcal{E}_L, \mathcal{E}_M) \approx \frac{(d!)^{1/d}\Gamma(1+1/d)p}{(|I_1| \cdots |I_{d-1}| \cdot p)^{1/d}}.$$

This is because $|M|$ is strongly concentrated around the mean of $B(I_1| \cdots |I_{d-1}|, p)$.

(h): The $|I_j|$ are strongly concentrated about their means which are of value $\approx np$. This results in replacing (15) by

$$E(W_n \mid \mathcal{E}_M) \approx \frac{(d!)^{1/d}\Gamma(1+1/d)}{n^{(d-1)/d}}.$$

Multiplying by n gives us part (a) of Theorem 5 with $c_d = (d!)^{1/d}\Gamma(1+1/d)$.

(i): The essential part of (b) is the inequality (21). We iteratively define X_m by $X_0 = \emptyset$ and for $m \geq 1$,

$$X_m = \left\{ (i_2, i_3, \ldots, i_d) \in \prod_{j=2}^{d} I_j \setminus \bigcup_{i<m} X_i : \right\}.$$

Here X_m defines that part of the data that is conditioned by the previous stps of the algorithm. Then we let

$$Y_m = (I_2 \times I_3 \times \cdots \times I_d) \cap \bigcup_{i<m} X_i.$$

We have to argue as in (18) that

$$\mathbf{Pr}(|Y_m| = O(Ln^{1/d})) = 1 - O(n^{-3}).$$

For this, where $S = \{x_{\mathbf{i}_t} : t = 1, 2, \ldots, s\}$ is a matching in H and $m \leq n - n^{1/(d+1)}$,

$$\mathbf{Pr}(S \subseteq Y_m) \leq \sum_{i_1 \leq \cdots \leq i_s} \mathbf{Pr}\left(\bigcap_{t=1}^{s} \{x_{\mathbf{i_t}} \in X_{i_t}\}\right) =$$

$$\sum_{i_1 \leq \cdots \leq i_s} \prod_{t=1}^{s} \mathbf{Pr}\left(x_{\mathbf{i_t}} \in X_{i_t} \middle| \bigcap_{\tau=1}^{t-1} \{x_{\boldsymbol{\tau}} \in X_{i_\tau}\}\right) \leq \sum_{i_1 \leq \cdots \leq i_s} \prod_{t=1}^{s} ((1+o(1))(n - i_t + 1)^{-(d-1)^2/d})$$

$$\leq \prod_{t=1}^{s} \sum_{l=1}^{n-m+1} \frac{1+o(1)}{(n-l+1)^{(d-1)^2/d}} \leq O\left(\frac{1}{n^{(d-1)^2/(d(d+1))}}\right)^s = O(n^{-3}),$$

for $s \geq 3d(d+1)/(d-1)^2$.

We deduce from this that we can write $|Y_m| = O(|I_2| + \cdots + |I_d|) = O(n^{1/d})$ as required. We can then replace (22) by

$$(c_d + o(1)) \sum_{m=1}^{n-n^{1/(d+1)}} \frac{1}{(n-m+1)^{(d-1)/d}} + n^{1/(d+1)} \approx dc_d n^{1/d}.$$

The final $n - n^{1/(d+1)}$ steps cost at most d per step and this completes our sketch proof of Theorem 5.

4 Greedy for small positive integer weights

When $W_{i,j,k}$ is chosen uniformly from $[1, M = n^\alpha], 0 < \alpha < 1$ we

(a): Let Z_m denote the cost of the mth triple. Then for $1 \leq m \leq n$ and $a \geq 1$,

$$\mathbf{Pr}(\exists m : Z_m \geq a) \leq n\left(1 - \frac{a}{M}\right)^{(n-m+1)^2} \leq n \exp\left\{-\frac{a(n-m+1)^2}{M}\right\} \leq n^{-2},$$

if

$$a \geq \frac{3M \log n}{(n-m+1)^2}. \tag{25}$$

Putting $m_0 = n - (3M \log n)^{1/2}$ we see that a satisfies (25) for

$$a = \begin{cases} 1 & m \leq m_0. \\ \left\lceil \frac{3M \log n}{(n-m+1)^2} \right\rceil & m > m_0. \end{cases}$$

It follows that w.h.p. and in expectation that if $m_1 = n - \log n$, then

$$\sum_{m=1}^{n} Z_m \leq m_0 + \sum_{m=m_0+1}^{m_1} \frac{3M \log n}{(n-m+1)^2} + M(n - m_1) = n + o(n),$$

5 Greedy versus Greedy

There is another version of the greedy algorithm where at each step we choose the “tple” of minimum weight that can be added to the current choice. Let $E(\lambda)$ denote the exponential rate λ random variable i.e. $\mathbf{Pr}(E(\lambda) \geq u) = e^{-\lambda u}$. We consider the d-dimensional case and argue next that if the weights $W_{i_1,\dots,i_d}$ are independent $E(1)$ then the value of the solution given by the two algorithms is the same in distribution. So let $G_{n,1}$ be the value returned by Algorithm 1 of Section 1 and let $G_{n,2}$ be the value returned by algorithm described in this section. We claim that $G_{n,1}$ and $G_{n,2}$ have the same distribution.

The distribution of $G_{n,1}$ is $E(n^{d-1}) + G_{n-1,1}$ and the distribution of $G_{n,2}$ is $E(n^d)(1 + (n - 1)) + G_{n-1,2}$. The term $E(n^d)(n-1)$ is a result of the fact that conditioning an exponential to be greater than x is equivalent to adding x to a copy of that variable. Then observe that the random variables $E(n^{d-1})$ and $nE(n^d)$ have the same distribution. The claim follows by induction.

Note that coincidentally, when $d = 3$, $\mathbf{E}(G_{n,1})$ is equal to the expected optimum value for the $d = 2$ case, see [5] and [6]. This does not generalise.

6 Final Comments

We have analysed a random multi-dimensional assignment problem with a particular form of objective fucntion. We have shown that w.h.p. there is a simple greedy algorithm that is a $(3 + o(1)$-approximation to the minimum. It is possible to replace the 3 here by $3 - \varepsilon$, by arguing that w.h.p. the optimum solution must use the (at least) second smallest j, k (when $d = 3$) for $\Omega(n)$ values of i. We omit the details as the real aim is to replace 3 by 1.

References

[1] K. Frankston, J. Kahn, B. Narayana and J. Park, Thresholds versus fractional expectation-thresholds.

[2] A.M. Frieze and G. Sorkin, Efficient algorithms for three-dimensional axial and planar random assignment problems, *Random Structures and Algorithms* 46 (2015) 160-196.

[3] A.M. Frieze and J. Yadegar, An algorithm for solving 3-dimensional assignment problems with application to scheduling a teaching practice, *Journal of the Operational Research Society* 32 (1981) 989-995.

[4] V. Kravtsov, Polynomial algorithms for finding the asymptotically optimal plan of the multiindex assignment problem, *Cybernetics and Systems Analysis* 41 (2005) 940-944.

[5] C. Nair, B. Prabhakar and M. Sharma, Proofs of the Parisi and Coppersmith-Sorkin random assignment conjectures, *Random Structures and Algorithms* 27 (2005) 413-444.

[6] S. Linusson and J. Wästlund, A proof of Parisi's conjecture on the random assignment problem, *Probability Theory and Related Fields* 128 (2004) 419-440.