How the Landscape of Random Job Shop Scheduling Instances Depends on the Ratio of Jobs to Machines

Matthew J. Streeter **Stephen F. Smith**

August 19, 2005
CMU-CS-05-162

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

We characterize the search landscape of random instances of the job shop scheduling problem (JSSP). Specifically, we investigate how the expected values of (1) backbone size, (2) distance between near-optimal schedules, and (3) makespan of random schedules vary as a function of the job to machine ratio ($\frac{N}{M}$). For the limiting cases $\frac{N}{M} \to 0$ and $\frac{N}{M} \to \infty$ we provide analytical results, while for intermediate values of $\frac{N}{M}$ we perform experiments. We prove that as $\frac{N}{M} \to 0$, backbone size approaches 100%, while as $\frac{N}{M} \to \infty$ the backbone vanishes. In the process we show that as $\frac{N}{M} \to 0$ (resp. $\frac{N}{M} \to \infty$), simple priority rules almost surely generate an optimal schedule, suggesting a theoretical account of the "easy-hard-easy" pattern of typical-case instance difficulty in job shop scheduling. We also draw connections between our theoretical results and the "big valley" picture of JSSP landscapes.

Keywords: job shop scheduling, big valley, easy-hard-easy pattern

1 Introduction

1.1 Motivations

The goal of this work is to provide a picture of the typical landscape of random instances of the job shop scheduling problem (JSSP), and to determine how this picture changes as a function of the job to machine ratio ($\frac{N}{M}$). Such a picture is potentially useful in (1) understanding how typical-case instance difficulty varies as a function of $\frac{N}{M}$ and (2) designing search heuristics that take advantage of regularities in typical instances of the JSSP.

1.1.1 Understanding instance difficulty as a function of $\frac{N}{M}$

In the job shop scheduling literature there is a conventional wisdom that square JSSPs (those with $\frac{N}{M} = 1$) are more difficult to solve than rectangular instances [7]. This work makes both theoretical and empirical contributions toward understanding this phenomenon. Empirically, we show that random schedules (resp. random local optima) are furthest from optimality when $\frac{N}{M} \approxeq 1$. Analytically, we prove that in the two limiting cases ($\frac{N}{M} \rightarrow 0$ and $\frac{N}{M} \rightarrow \infty$) there exist simple priority rules that almost surely produce an optimal schedule, providing theoretical evidence of an "easy-hard-easy" pattern of instance difficulty in the JSSP.

1.1.2 Informing the design of search heuristics

Heuristics based on local search (e.g., taboo search [8, 18], iterated local search [12]) have shown excellent performance on benchmark instances of the job shop scheduling problem [9, 10]. In order to design an effective heuristic, one must (explicitly or implicitly) make assumptions about the search landscape of instances to which the heuristic will be applied. For example, Nowicki and Smutnicki motivate the use of *path relinking* in their state-of-the-art *i*-TSAB algorithm by citing evidence that the JSSP has a "big valley" distribution of local optima [20]. One of the conclusions of our work is that the typical landscape of random instances can only be thought of as a big valley for values of $\frac{N}{M}$ close to 1; for larger values of $\frac{N}{M}$ (including values common in benchmark instances), the landscape breaks into many big valleys, suggesting that modifications to *i*-TSAB may allow it to better handle this case.

1.2 Contributions

The contributions of this paper are twofold. First, we design a novel set of experiments and run these experiments on random instances of the JSSP. Second, we derive analytical results that confirm and provide insight into the trends suggested by our experiments.

The main contributions of our empirical work are as follows.

- For low values of $\frac{N}{M}$, we show that low-makespan schedules are clustered in a small region of the search space and many attributes (i.e., directed disjunctive graph edges) are common to all low-makespan schedules. As $\frac{N}{M}$ increases, low-makespan schedules become dispersed throughout the search space and there are no attributes common to all low-makespan schedules.

- We introduce a statistic (neighborhood exactness) that can be used to quantitatively measure the "ruggedness" of a search landscape, and estimate the expected value of this statistic for random instances of the JSSP. These results, in combination with the results on clustering, suggest that the landscape of typical instances of the JSSP can be described as a big valley only for low values of $\frac{N}{M}$; for high values of $\frac{N}{M}$ there are many separate big valleys.

For the limiting cases $\frac{N}{M} \to 0$ and $\frac{N}{M} \to \infty$, we derive analytical results. Specifically, we prove that

- as $\frac{N}{M} \to 0$, the expected size of the backbone (i.e., the set of problem variables that have a common value in all global optima) approaches 100%, while as $\frac{N}{M} \to \infty$, the expected backbone size approaches 0%; and
- as $\frac{N}{M} \to 0$ (resp. $\frac{N}{M} \to \infty$), a randomly-generated schedule will almost surely have a makespan that is near-optimal, and will be located "close" (in a sense to be precisely defined) to an optimal schedule.

2 Related Work

There are at least three threads of research that have conducted search space analyses related to the ones we conduct here. These include literature on the "big valley" distribution common to a number of combinatorial optimization problems, studies of backbone size in boolean satisfiability, and a statistical mechanical analysis of the TSP. We briefly review these three areas below, as well as relevant work on phase transitions and the "easy-hard-easy" pattern of instance difficulty.

2.1 The Big Valley

The term "big valley" originated in a paper by Boese et al. [4] that examined the distribution of local optima in the Traveling Salesman Problem (TSP). Based on a sample of local optima obtained by next-descent starting from random TSP tours, Boese calculated two correlations:

1. the correlation between the cost of a locally optimal tour and its average distance to other locally optimal tours, and
2. the correlation between the cost of a locally optimal tour and the distance from that tour to the best tour in the sample.

The distance between two TSP tours was defined as the total number of edges minus the number of edges that are common to the two tours. Based on the fact that both of these correlations were surprisingly high, and the fact that the mean distance between random local optima was small relative to the mean distance between random tours, Boese conjectured that local optima in the

TSP are arranged in a "big valley". The term does not have a formal definition, but intuitively it means that if the search space could be locally smoothed in some manner it would then have a single basin of attraction with respect to common TSP move operators such as Lin 2-opt.

Boese's analysis has been applied to other combinatorial problems [11], including the permutation flow shop scheduling problem [28, 21] and the JSSP [19]. Correlations observed for the JSSP are generally weaker than those observed for the TSP.

2.2 Backbone Size

The *backbone* of a problem instance is the set of attributes common to all globally optimal solutions of that instance. For example, in the boolean satisfiability problem (SAT), the backbone is the set of variable assignments that are common to all satisfying assignments. In the JSSP, the backbone has been defined as the number of disjunctive edges (§3.2) that have a common orientation in all globally optimal schedules (a formal definition is given in §4).

There is a large literature on backbones in combinatorial optimization problems, including many empirical and analytical results [23, 17]. In an analysis of problem difficulty in the JSSP, Watson et al. [29] present histograms of backbone size for random 6x6 (6 job, 6 machine) and 6x4 (6 job, 4 machine) JSSP instances. Summarizing experiments not reported in their paper, Watson et al. note that "For [job:machine ratios] > 1.5, the bias toward small backbones becomes more pronounced, while for ratios < 1, the bias toward larger backbones is further magnified." §4 generalizes these observations and proves two theorems that give insight into why this phenomenon occurs.

2.3 Statistical Mechanical Analyses

A large and growing literature applies techniques from statistical mechanics to the analysis of combinatorial optimization problems [14]. At least one result obtained in this literature concerns clustering of low-cost solutions. In a study of the TSP, Mézard and Parisi [16] obtain an expression for the expected overlap (number of common edges) between random TSP tours drawn from a Boltzmann distribution. They show that as the temperature parameter of the Boltzmann distribution is lowered (placing more probability mass on low-cost TSP tours), expected overlap approaches 100%. Though we do not use a Boltzmann weighting, §5 of this paper examines how expected overlap between random JSSP schedules changes as more probability mass is placed on low-makespan schedules.

2.4 Phase Transitions and the Easy-hard-easy Pattern

Loosely speaking, a phase transition occurs in a system when the expected value of some statistic varies discontinuously (asymptotically) as a function of some parameter. As an example, for any $\epsilon > 0$ it holds that random instances of the 2-SAT problem are satisfiable with probability asymptotically approaching 1 when the clause to variable ratio ($\frac{m}{n}$) is $1 - \epsilon$, but are satisfiable with probability approaching 0 when the clause to variable ratio is $1 + \epsilon$. A similar statement is

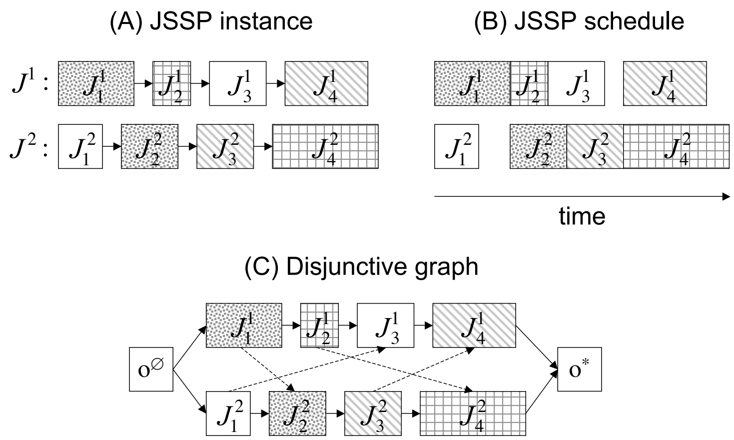

Figure 1: (A) A JSSP instance, (B) a feasible schedule for the instance, and (C) the disjunctive graph representation of the schedule. Boxes represent operations; operation durations are proportional to the width of a box; and the machine on which an operation is performed is represented by texture. In (C), solid arrows represent conjunctive arcs and dashed arrows represent disjunctive arcs.

conjectured to hold for 3-SAT; the critical value k of $\frac{m}{n}$ (if it exists) must satisfy $3.42 \leq k \leq 4.51$ [1].

For some problems that exhibit phase transitions (notably 3-SAT), average-case instance difficulty (for typical solvers) appears to first increase and then decrease as one increases the relevant parameter, with the hardest instances appearing close to the threshold value [6]. This phenomenon has been referred to as an "easy-hard-easy" pattern of instance difficulty [13]. In §7.4 we discuss evidence of an easy-hard-easy pattern of instance difficulty in the JSSP, though (to our knowledge) it is not associated with any phase transition.

3 The Job Shop Scheduling Problem

We adopt the notation $[n] \equiv \{1, 2, \ldots, n\}$.

3.1 Problem Definition

Definition (JSSP instance). *An* N *by* M JSSP instance $I_{N,M} = \{J^1, J^2, \ldots, J^N\}$ *is a set of* N jobs *, where each job* $J^k = (J^k_1, J^k_2, \ldots, J^k_M)$ *is a sequence of* M operations, *and each operation* J^k_i *is a triple* $(k, \bar{m}, \bar{\tau})$ *where* $\bar{m} \in [M]$ *is a* machine number *and* $\bar{\tau} > 0$ *is a* duration. *We require that each job use each machine exactly once (i.e., for each* $J^k \in I_{N,M}$ *and* $\bar{m} \in [M]$*, we have* $|\{o = (k, \bar{m}', \bar{\tau}) \in J : \bar{m}' = \bar{m}\}| = 1$*). We define*

1. $ops(I_{N,M}) \equiv \{o \in J : J \in I_{N,M}\}$,

2. $m((k,\bar{m},\bar{\tau})) \equiv \bar{m}$,

3. $\tau((k,\bar{m},\bar{\tau})) \equiv \bar{\tau}$,

4. $\tau(O) \equiv \sum_{o \in O} \tau(o)$ *(where O is any set of operations), and*

5. *the* job-predecessor $\mathcal{J}(J_i^k)$ *of an operation* J_i^k *as*

$$\mathcal{J}(J_i^k) \equiv \begin{cases} J_{i-1}^k & \textit{if } i > 1 \\ o^{\emptyset} & \textit{otherwise} \end{cases}$$

where $o^{\emptyset}$ *is a fictitious operation,* $\tau(o^{\emptyset}) = 0$*, and* $m(o^{\emptyset})$ *is undefined.*

Definition (JSSP schedule). *A JSSP schedule for an instance* $I_{N,M}$ *is a function* $S : ops(I_{N,M}) \rightarrow \Re_+$ *that associates with each operation* $o \in ops(I_{N,M})$ *a* start time $S(o)$. *We make the following definitions.*

1. *The* completion time *of an operation* o *is* $S^+(o) \equiv S(o) + \tau(o)$.

2. *Let* $\bar{O}(o) = \{\bar{o} \in ops(I_{N,M}) : m(\bar{o}) = m(o), S(\bar{o}) < S(o)\}$ *denote the set of operations scheduled to run before an operation* o *on* o*'s machine. The* machine-predecessor $\mathcal{M}(o)$ *of* o *is defined as*

$$\mathcal{M}(o) \equiv \begin{cases} \arg\max_{\bar{o} \in \bar{O}(o)} S(\bar{o}) & \textit{if } \bar{O}(o) \neq \emptyset \\ o^{\emptyset} & \textit{otherwise.} \end{cases}$$

3. S *is a* feasible *schedule if* $S(o) \geq \max(S^+(\mathcal{J}(o)), S^+(\mathcal{M}(o))) \; \forall o \in ops(I_{N,M})$.

4. *The quantity*

$$\ell(S) \equiv \max_{o \in ops(I_{N,M})} S^+(o)$$

is called the makespan *of* S.

We consider the makespan-minimization version of the JSSP, in which the goal is to find a schedule that minimizes the makespan.

For the remainder of the paper, whenever we refer to a JSSP schedule S we shall assume that

$$S(o) = \max(S^+(\mathcal{J}(o)), S^+(\mathcal{M}(o))) \; \forall o \in ops(I_{N,M}) \tag{3.1}$$

(i.e., S is a so-called *semi-active* schedule). In other words, we ignore schedules with superfluous idle time between the end of one operation and the start of another.

Figure 1 (A) and (B) depict, respectively, a JSSP instance and a feasible schedule for that instance.

3.2 Disjunctive Graphs

A schedule satisfying (3.1) can be uniquely represented by a weighted, directed graph called its *disjunctive graph*.

Definition (disjunctive graph). *The disjunctive graph* $G = G(I_{N,M}, S)$ *of a schedule* S *for a JSSP instance* $I_{N,M}$ *is the weighted, directed graph* $G = (V, \vec{E}, w)$ *defined as follows.*

- $V = ops(I_{N,M}) \cup \{o^{\emptyset}, o^*\}$*, where* $o^{\emptyset}$ *(resp.* o^**) is a fictitious operation with* $\tau(o^{\emptyset}) = 0$ *and* $m(o^{\emptyset})$ *undefined.*
- $\vec{E} = \vec{C} \cup \vec{D}$*, where*
  - $\vec{C} = \{(o_1, o_2) : \{o_1, o_2\} \subseteq ops(I_{N,M}), \mathcal{J}(o_2) = o_1\} \cup \left\{(o^{\emptyset}, J_1) : J \in I_{N,M}\right\}$ $\cup \left\{(J_M, o^*) : J \in I_{N,M}\right\}$*, and*
  - $\vec{D} = \{(o_1, o_2) : \{o_1, o_2\} \subseteq ops(I_{N,M}), m(o_1) = m(o_2), S(o_1) < S(o_2)\}$.
- $w((o_1, o_2)) = \tau(o_1)$.

The arcs in $\vec{C}$ *are called* conjunctive arcs, *while those in* $\vec{D}$ *are called* disjunctive arcs. *We make the following definitions*

1. *The set* $E(I_{N,M})$ *of* disjunctive edges *of* $I_{N,M}$ *is defined by*
$$E(I_{N,M}) \equiv \{\{o_1, o_2\} \subseteq ops(I_{N,M}) : m(o_1) = m(o_2)\}$$
2. *For a disjunctive edge* $e \in E(I_{N,M})$*, the function* $\vec{e}(S)$ *returns the unique arc* $a \in \{(o_1, o_2), (o_2, o_1)\}$ *which appears in* $G(I_{N,M}, S)$.
3. *The* disjunctive graph distance $\|S_1 - S_2\|$ *between two schedules* S_1 *and* S_2 *for* $I_{N,M}$ *is defined by*
$$\|S_1 - S_2\| \equiv |\{e \in E(I_{N,M}) : \vec{e}(S_1) \neq \vec{e}(S_2)\}|$$

It is straightforward to verify the following Proposition [22].

Proposition 1. *If* S *is a feasible schedule for* $I_{N,M}$ *satisfying assumption* (3.1)*, then* $\ell(S)$ *is equal to the length of the longest weighted path from* $o^{\emptyset}$ *to* o^* *in* $G(I_{N,M}, S)$.

Figure 1 (C) depicts the disjunctive graph for the schedule depicted in Figure 1 (B).

3.3 Random Schedules and Instances

We define a uniform distribution over JSSP instances as follows.

Definition (random JSSP instance). *A random N by M JSSP instance $I_{N,M}$ is generated as follows.*

1. *Let $\phi_1, \phi_2, \ldots, \phi_N$ be random permutations of $[M]$.*
2. *Let the elements of $\{\tau_{k,i} : k \in [N], i \in [M]\}$ be drawn independently at random from a common distribution over $(0, \tau_{max}]$ with mean μ variance $\sigma^2 > 0$.*
3. *Define $I_{N,M} = \{J^1, J^2, \ldots, J^N\}$, where $J_i^k = (\phi_k(i), \tau_{k,i})$.*

Note that we assume a maximum operation duration τ_{max}. We choose operation durations from a uniform distribution over $\{1, 2, \ldots, 100\}$ for the experiments described in this paper.

To define a distribution over JSSP schedules, it is convenient to first make the following definition.

Definition (priority rule). *A* priority rule π *is a function that, given an instance $I_{N,M}$, returns a sequence $T = \pi(I_{N,M})$, where the set of elements of T is $ops(I_{N,M})$, and where, for any $\{J_i^k, J_j^k\} \subseteq ops(I_{N,M})$ with $i < j$, J_i^k appears before J_j^k in T.*

The schedule $S = S(\pi, I_{N,M})$ associated with π is defined by the following procedure.

1. *Set $T := \pi(I_{N,M})$. Let T_i denote the i^{th} operation in T.*
2. *Set $S(o) := \infty \; \forall o \in ops(I_{N,M})$ and set $S(o^\emptyset) := 0$.*
3. *For i from 1 to NM do:*
   (a) *Set $S(T_i) := \max(S^+(\mathcal{J}(T_i)), S^+(\mathcal{M}(T_i)))$.*

A priority rule is called instance-independent *if $T = \pi(I_{N,M})$ depends only on N and M.*

Our definition of random schedules is equivalent to the one proposed by Mattfeld [15].

Definition (random schedule). *A random schedule for an N by M instance $I_{N,M}$ is a schedule obtained using the priority rule π_{rand}, where $\pi_{rand}(I_{N,M})$ returns a random element of the set $\{T = \pi(I_{N,M}) : \pi$ is a priority rule$\}$.*

4 Number of Common Attributes as a Function of Makespan

In this section we compute the expected value of $|\rho\_backbone|$ (defined below) as a function of ρ for random *N* by *M* JSSP instances, and examine how the shape of this curve changes as a function of $\frac{N}{M}$. We make the following definition (a related definition appears in [23]).

Definition ($\rho\_backbone$). *Let $\hat{S}$ be an optimal schedule for a JSSP instance $I_{N,M}$. Let $\rho\_opt(I_{N,M}) \equiv \{S : \ell(S) \leq (1+\rho)\ell(\hat{S})\}$ be the set of schedules whose makespan is within a factor ρ of optimality. Then*

$$\rho\_backbone(I_{N,M}) \equiv \{e \in E(I_{N,M}) : \vec{e}(S) = \vec{e}(S') \; \forall \{S, S'\} \subseteq \rho\_opt(I_{N,M})\} \, .$$

4.1 Computing $|\rho\_backbone|$

For an arbitrary JSSP instance, we compute $|\rho\_backbone|$ for all ρ using multiple runs of a branch and bound algorithm.

Let $opt(a)$ denote the minimum makespan among schedules whose disjunctive graphs contain the arc a. In branch and bound algorithms for the JSSP, nodes in the search tree represent choices of orientations for a subset of the disjunctive edges. Thus, by constructing a root search tree node that has a as a fixed arc, we can determine $opt(a)$ using existing branch and bound algorithms. We use an algorithm due to Brucker et al. [5] because it is efficient and because the code for it is freely available via ORSEP.

Suppose we are interested in all schedules whose makespan is at most $t = \rho \cdot \hat{\ell}$, where $\hat{\ell}$ is the optimum makespan for the given instance. Consider an arbirary disjunctive edge $e = \{o_1, o_2\}$, with orientations $a_1 = (o_1, o_2)$ and $a_2 = (o_2, o_1)$. It must be the case that $\min(opt(a_1), opt(a_2)) = \hat{\ell}$. If $t < \max(opt(a_1), opt(a_2))$, then one of the two orientations of e precludes schedules with makespan $\leq t$, so e must belong to the $\rho\_backbone$. Thus

$$|\rho\_backbone| = \sum_{\{o_1,o_2\} \in E} [t < \max(opt((o_1, o_2)), opt((o_2, o_1))] \tag{4.1}$$

where the [...] notation indicates a function that returns 1 if the predicate enclosed in the brackets is true, 0 otherwise.

Using (4.1), we can determine $|\rho\_backbone|$ for all ρ by performing $1 + M\binom{N}{2}$ runs of branch and bound. The first branch and bound run is used to find a globally optimal schedule, which gives the value of opt for one of the two possible orientations of each of the $M\binom{N}{2}$ disjunctive edges. A separate branch and bound run is used to determine the values of opt for the $M\binom{N}{2}$ alternative orientations.

Figure 2 graphs the fraction of disjunctive edges that belong to the $\rho\_backbone$ as a function of ρ for instance ft10 (a 10 job, 10 machine instance) from the OR library [3]. Note that by definition the curve is non-increasing with respect to ρ, and that the curve is exact for all ρ. It is noteworthy that among schedules within 0.5% of optimality, 80% of the disjunctive edges have a fixed orientation. We will see that this behavior is typical of JSSP instances with $\frac{N}{M} = 1$.

4.2 Results

We plotted $|\rho\_backbone|$ as a function of ρ for all instances in the OR library having 10 or fewer jobs and 10 or fewer machines. The results are available online [25]. Inspection of the graphs revealed that the shape of the curve is largely a function of the job:machine ratio. To investigate this further, we repeated these experiments on a large number of randomly generated JSSP instances.

We use randomly-generated instances with 7 different combinations of N and *M* to study instances with $\frac{N}{M}$ equal to 1, 2, or 3. For $\frac{N}{M} = 1$ we use 6x6, 7x7, and 8x8 instances; for $\frac{N}{M} = 2$ we use 8x4 and 10x5 instances; and for $\frac{N}{M} = 3$ we use 9x3 and 12x4 instances. We generate 1000 random instances for each combination of N and M.

Figure 3 presents the expected fraction of edges belonging to the ρ-backbone as a function of ρ for each combination of N and M, grouped according to $\frac{N}{M}$. For the purposes of this study the

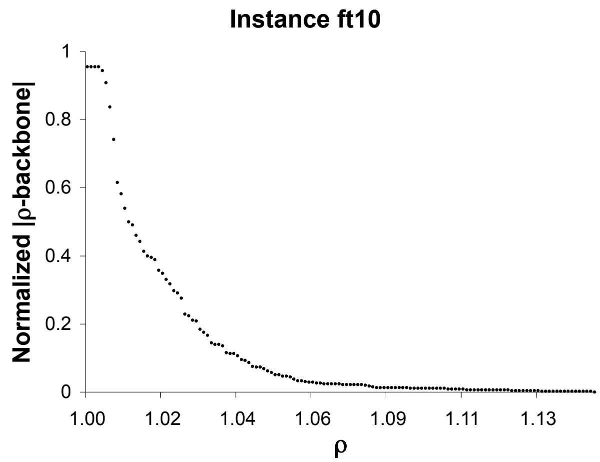

Figure 2: Normalized $|\rho\_backbone|$ as a function of ρ for OR library instance ft10.

two most important observations about Figure 3 are as follows.

- The curves depend on both the size of the instance (i.e., NM) and the shape (i.e., $\frac{N}{M}$). Of these two factors, $\frac{N}{M}$ has by far the stronger influence on the shape of the curves.
- For *all* values of ρ, the expected fraction of edges belonging to the $\rho\_backbone$ decreases as $\frac{N}{M}$ increases.

4.3 Analysis

We now give some insight into Figure 3 by analyzing two limiting cases. We prove that as $\frac{N}{M}\rightarrow 0$, the expected fraction of disjunctive edges that belong to the backbone approaches 1, while as $\frac{N}{M}\rightarrow\infty$ this expected fraction approaches 0.

Intuitively, what happens is as follows. As $\frac{N}{M}\rightarrow 0$ (i.e., N is held constant and $M\rightarrow\infty$) each of the jobs becomes very long. Individual disjunctive edges then represent precedence relations among operations that should be performed very far apart in time. For example, if there are 10,000 machines (and so each job consists of 10,000 operations), a disjunctive edge might specify whether operation 1,200 of job A is to be performed before operation 8,500 of job B. Clearly, waiting for job B to complete 8,500 of its operations before allowing job A to complete 12% of its operations is likely to produce an inefficient schedule. Thus, orienting a single disjunctive edge in the "wrong" direction is likely to prevent a schedule from being optimal, and so any particular edge will likely have a common orientation in all globally optimal schedules.

In contrast, when $\frac{N}{M}\rightarrow\infty$, it is the workloads of the machines that become very long. The order in which the jobs are processed on a particular machine does not matter much as long as the machine with the longest workload is kept busy, and so the fact that a particular edge is oriented a particular way is unlikely to prevent a schedule from being optimal. All of this is formalized below.

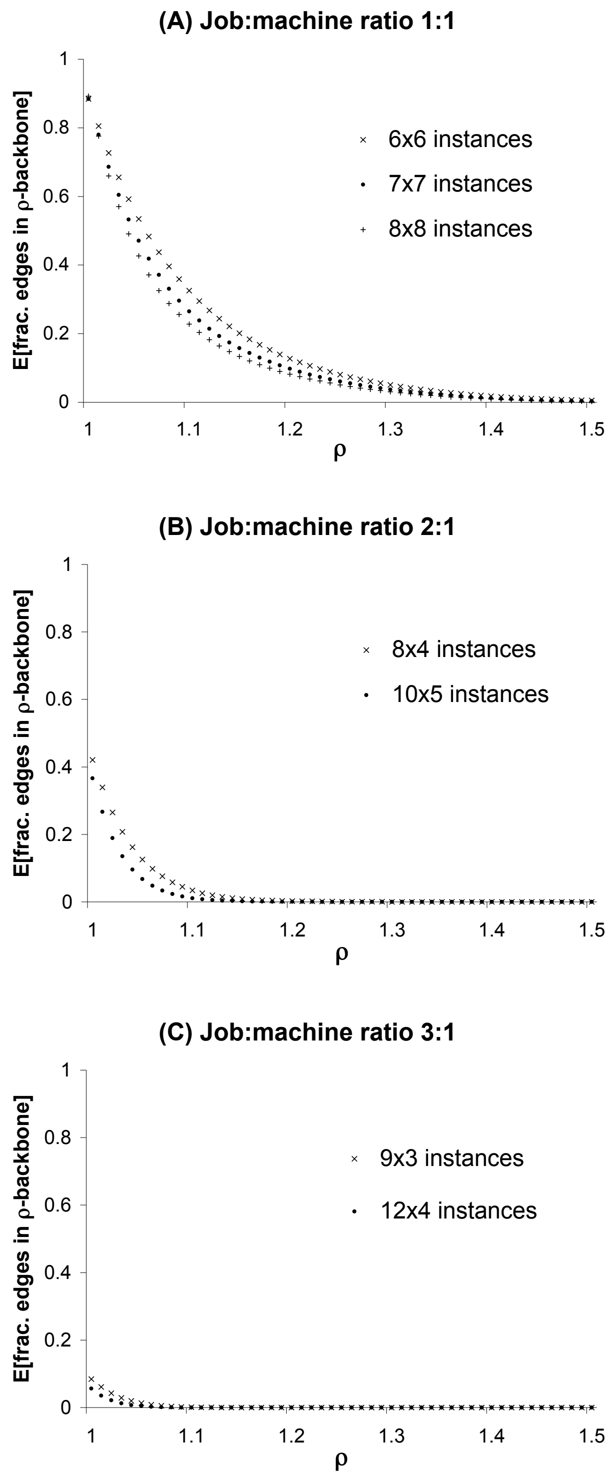

Figure 3: Expected fraction of edges in ρ-backbone as a function of ρ for random JSSP instances. Graphs (A), (B), and (C) depict curves for random instances with $\frac{N}{M}$ = 1, 2, and 3, respectively.

We will make use of the following well-known definition.

Definition (whp). *A sequence of events ξ_n occurs with high probability (whp) if $\lim_{n\to\infty}\mathbb{P}[\xi_n]=1$.*

Lemma 1 and Theorem 1 show that for constant N, a randomly chosen edge of a random N by M JSSP instance will be in the backbone whp (as $M\to\infty$). Lemma 2 and Theorem 2 show that for constant M, a randomly chosen edge of a random N by M JSSP instance will not be in the backbone whp (as $N\to\infty$). The proof of Lemma 2 appears in Appendix A.

Lemma 1. *Let $I_{N,M}$ be a random N by M JSSP instance, and let $S=S(\pi,I_{N,M})$ be the schedule for $I_{N,M}$ obtained using some instance-independent priority rule π. For an arbitrary job $J\in I_{N,M}$, let $\Delta_J^S\equiv S^+(J_M)-\tau(J)$ be the amount of time by which the completion of J is delayed due to resource constraints. Then $\mathbb{E}[\Delta_J]$ is $O(N)$.*

Proof. We assume $N=2$ and $M>1$. The generalization to larger N is straightforward, while the cases $N=1$ and $M=1$ are trivial. Let $I_{N,M}=\{J^1,J^2\}$ and let $J=J^1$. We say that an operation J_i^1 *overlaps* with an operation J_j^2 if

1. J_j^2 appears before J_i^1 in $\pi(I_{N,M})$, and
2. $[S(J_j^2),S^+(J_j^2)]\cap[S^+(J_{i-1}^1),S^+(J_{i-1}^1)+\tau(J_i^1)]\neq\emptyset$.

If additionally $m(J_i^1)=m(J_j^2)$, we say that J_i^1 *contends* with J_j^2. Let $\theta_{i,j}$ (resp. $\delta_{i,j}$) be an indicator for the event that J_i^1 overlaps (resp. contends) with J_j^2. Let $C_i\equiv\{J_j^2:\theta_{i,j}=1\}$ be the set of operations in J^2 that J_i^1 overlaps with. Then $|C_i\cap\bigcup_{i'>i}C_{i'}|\le 1$. Thus

$$\sum_i|C_i|=\sum_i|C_i\setminus\bigcup_{i'>i}C_{i'}|+\sum_i|C_i\cap\bigcup_{i'>i}C_{i'}|\le 2M\ . \tag{4.2}$$

Let $\bar{I}=I_{N,M-1}$ be a random N by $M-1$ JSSP instance, and define $\bar{\theta}_{i,j}$, $\bar{\delta}_{i,j}$, and $\bar{C}_i$ analogously to the above. Then for $i,j\le M-1$,

$$\mathbb{P}\left[\theta_{i,j}=1|m(J_i^1)=m(J_j^2)\right]=\mathbb{P}\left[\bar{\theta}_{i,j}=1\right]\ .$$

This is true because $\mathbb{P}[\theta_{i,j}=1]$ is a function of the joint distribution of the operations in the set $\{J_{i'}^1:i'<i\}\cup\{J_{j'}^2:j'<j\}$; and, as far as this joint distribution is concerned, conditioning on the event $m(J_i^1)=m(J_j^2)$ is like deleting the operations that use the machine $m(J_i^1)$.

Thus $\mathbb{E}\left[\delta_{i,j}\right]=\frac{1}{M}\mathbb{E}\left[\theta_{i,j}|m(J_i^1)=m(J_j^2)\right]=\frac{1}{M}\mathbb{E}\left[\bar{\theta}_{i,j}\right]$. Therefore,

$$\begin{aligned}\textstyle\sum_{i=1}^M\sum_{j=1}^M\mathbb{E}[\delta_{i,j}] &\le 2+\textstyle\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\mathbb{E}[\delta_{i,j}]\\ &=2+\tfrac{1}{M}\textstyle\sum_{i=1}^{M-1}\sum_{j=1}^{M-1}\mathbb{E}[\bar{\theta}_{i,j}]\\ &=2+\tfrac{1}{M}\textstyle\sum_{i=1}^{M-1}\mathbb{E}[|\bar{C}_i|]\\ &\le 4\end{aligned}$$

where in the last step we have used (4.2). It follows that $\mathbb{E}[\Delta_J^S]\le 4\tau_{max}$ (τ_{max} is the maximum operation duration defined in §3). When we consider arbitrary N, we get $\mathbb{E}[\Delta_J^S]\le 4\tau_{max}(N-1)$. $\square$

Corollary 1. *Let $I_{N,M}$ be a random N by M JSSP instance. Then for fixed N, it holds whp (as $M \to \infty$) that an optimal schedule $\hat{S}$ for $I_{N,M}$ has $\ell(\hat{S}) = \max_{k \in [N]} \tau(J^k)$.*

Proof. Let π_k be an instance-independent priority rule that first schedules all the operations in the job J^k, then schedules the remaining operations in some arbitrary order. Let π_{max} be a priority rule (not instance-independent) that first computes $k^* = \arg\max_{k \in [N]} \tau(J^k)$, and then invokes π_{k^*}. Using Lemma 1, we have

$$\mathbb{E}[\Delta_J^{\pi_{max}}] \leq \sum_k \mathbb{E}[\Delta_J^{\pi_k}] = O(N^2)$$

where $J \in I_{N,M}$ and we define $\Delta_J^{\pi} \equiv \Delta_J^{S(\pi, I_{N,M})}$. Then by Markov's inequality, $\Delta_J^{\pi_{max}} < M^{\frac{1}{4}}$ $\forall J \in I_{N,M}$ whp. By the Central Limit Theorem, each $\tau(J)$ is asymptotically normally distributed with mean μM and standard deviation $\sigma\sqrt{M}$. It follows that whp, $\tau(J^{k^*}) - \tau(J^k) > M^{\frac{1}{4}}$ $\forall k \neq k^*$. This implies $\ell(\hat{S}) \leq \ell(S(\pi_{max}, I_{N,M})) = \tau(J^{k^*})$. □

Theorem 1. *Let $I_{N,M}$ be a random N by M JSSP instance, and let e be a randomly selected element of $E(I_{N,M})$. Then for fixed N, it holds whp (as $M \to \infty$) that $e \in 1\_backbone(I_{N,M})$.*

Proof. Let $e = \{J_i, J'_j\}$ with $i \leq j$, let $a = (J'_j, J_i)$, and let $S_a = \{S : \vec{e}(S) = a\}$. We show that whp, $\min_{S \in S_a} \ell(S) > \max_{J \in I_{N,M}} \tau(J)$. By Corollary 1, this implies $e \in 1\_backbone(I_{N,M})$ whp.

Assume $j - i \geq M^{\frac{3}{4}}$ (holds whp). Let $\hat{S} = \arg\min_{S \in S_a} \ell(S)$. Then

$$P = (o^{\emptyset}, J'_1, J'_2, \ldots, J'_j, J_i, J_{i+1}, \ldots, J_M, o^*)$$

is a directed path in $G(I_{N,M}, \hat{S})$ that passes through at least $3 + M + M^{\frac{3}{4}}$ vertices. By Proposition 1, $\ell(\hat{S}) \geq w(P)$, where $w(P)$ is the weighted length of P. It remains to show that $w(P) > \max_{J \in I_{N,M}} \tau(J)$ whp. For any fixed P, it follows by the Central Limit Theorem that $w(P)$ is asymptotically normally distributed with mean $\mu(|P| - 2)$ and standard deviation $\sigma\sqrt{(|P| - 2)}$, while for each J, $\tau(J)$ is asymptotically normally distributed with mean μM and standard deviation $\sigma\sqrt{M}$. That $w(P) > \max_{J \in I_{N,M}} \tau(J)$ whp follows by Chebyshev's inequality.

□

Lemma 2. *Let $I_{N,M}$ be a random N by M JSSP instance. Let ϕ be an arbitrary permutation of $[N]$, and consider the priority rule π defined by*

$$\pi(I_{N,M}) = T_1 T_2 \ldots T_M$$

where

$$T_i = \left(J_i^{\phi(1)}, J_i^{\phi(2)}, \ldots, J_i^{\phi(N)}\right).$$

Then for fixed M, there exists a $K > 0$ such that with probability $1 - o(\exp(-KN))$, the schedule $\hat{S} = S(\pi, I_{N,M})$ has the property that

$$\hat{S}(o) = \hat{S}^+(\mathcal{M}(o)) \; \forall o \in ops(I_{N,M}).$$

Proof. See Appendix A. □

Corollary 2. *Let $I_{N,M}$ be a random N by M JSSP instance, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$. Let $\tau(\bar{m}) \equiv \tau(\{o \in ops(I_{N,M}) : m(o) = \bar{m}\})$ denote the workload of machine $\bar{m}$. Then for fixed M, it holds whp (as $N \to \infty$) that $\ell(\hat{S}) = \max_{\bar{m} \in [M]} \tau(\bar{m})$.*

Corollary 2 confirms a conjecture of Taillard [26].

Theorem 2. *Let $I_{N,M}$ be a random N by M JSSP instance, and let e be a randomly selected element of $E(I_{N,M})$. Then for fixed M, it holds whp (as $N \to \infty$) that $e \notin 1\_backbone(I_{N,M})$.*

Proof. Let $e = \{J_i, J'_j\}$. Remove both J and J' from $I_{N,M}$ to create a new instance $I_{N-2,M}$, which comes from the same distribution as a random $N-2$ by M JSSP instance. Lemma 2 shows that whp there exists an optimal schedule $\hat{S}$ for $I_{N-2,M}$ with the property described in the statement of the lemma.

Let $\tau(\bar{m}) \equiv \tau(\{o \in ops(I_{N,M}) : m(o) = \bar{m}\})$ denote the workload of machine $\bar{m}$. By the Central Limit Theorem, each $\tau(\bar{m})$ is asymptotically normally distributed with mean $\mu(N-2)$ and standard deviation $\sigma\sqrt{N-2}$. It follows that whp, $|\tau(\bar{m}) - \tau(\bar{m}')| > N^{\frac{1}{4}}\ \forall \bar{m} \neq \bar{m}'$.

Thus whp there will be only one machine still processing operations during the interval $[\ell(\hat{S}) - N^{\frac{1}{4}}, \ell(\hat{S})]$. Because $\max(\tau(J), \tau(J')) \leq M\tau_{max} = O(1)$, we can use this interval to construct optimal schedules containing the disjunctive arc (J_i, J'_j) as well as optimal schedules containing the disjunctive arc (J'_j, J_i). □

5 Clustering as a Function of Makespan

In this section we estimate the expected distance between random schedules within a factor ρ of optimality, as a function of ρ for various combinations of N and M. We examine how the shape of this curve changes as a function of $\frac{N}{M}$. More formally, if

- $I_{N,M}$ is a random N by M JSSP instance,
- $\hat{S}$ is an optimal schedule for $I_{N,M}$,
- $\rho\_opt(I_{N,M}) \equiv \{S : \ell(S) \leq (1+\rho)\ell(\hat{S})\}$, and
- S_1^ρ and S_2^ρ are drawn independently at random from $\rho\_opt(I_{N,M})$,

we wish to compute $\mathbb{E}[||S_1^\rho - S_2^\rho||]$.

The experiments of the previous section provide an upper bound on this quantity:

$$\mathbb{E}\left[||S_1^\rho - S_2^\rho||\right] \leq M\binom{N}{2} - \mathbb{E}\left[|\rho\_backbone|\right]$$

but provide no lower bound.

5.1 Methodology

We generate "random" samples from $\rho\_opt(I_{N,M})$ by running the simulated annealing algorithm of van Laarhoven et al. [27] until it finds such a schedule. More precisely, our procedure for sampling distances is as follows.

1. Generate a random N by M JSSP instance I.
2. Using the branch and bound algorithm of Brucker et al. [5], determine the optimal makespan of I.
3. Perform two runs, R_1 and R_2, of the van Laarhoven et al. [27] simulated annealing algorithm. Restart each run as many times as necessary for it to find a schedule whose makespan is optimal.
4. For each $\rho \in \{1, 1.01, 1.02, ..., 2\}$, find the first schedule, call it $s_i(\rho)$, in each run R_i whose makespan is within a factor ρ of optimality. Add the distance between $s_1(\rho)$ and $s_2(\rho)$ to the sample of distances associated with ρ.

Note that the distances for different values of ρ are dependent, but that for a given ρ all the sampled distances are independent, so that our estimates are unbiased. Figure 4 presents the results of running this procedure on 1000 random JSSP instances for the same 7 combinations of N and M that were used in §4.2.

From examination of Figure 4, we see that for ρ near 1, the ρ-optimal schedules are in fact dispersed widely throughout the search space for $\frac{N}{M} = 3$, and that this is true to a lesser extent for $\frac{N}{M} = 2$.

An immediate implication of Figure 4 is that whether or not they exhibit the two correlations that are the operational definition of a big valley, typical landscapes for JSSP instances with $\frac{N}{M} = 3$ cannot be expected to be big valleys in the intuitive sense of these words. If anything, one might posit the existence of multiple big valleys, each leading to a separate global optimum. The next section expands upon these observations.

6 The Big Valley

In the section we formalize the notion of a big valley landscape, conduct experiments to determine the extent to which random JSSP instances exhibit such a landscape as we vary $\frac{N}{M}$, and present analytical results for the limiting cases $\frac{N}{M} \to 0$ and $\frac{N}{M} \to \infty$.

6.1 Formalization

The following three definitions allow us to formalize the notion of a big valley landscape.

Definition (Neighborhood $\mathcal{N}_r$). *Let I be an arbitrary JSSP instance, and let U be the set of all schedules for I. The neighborhood $\mathcal{N}_r : U \to 2^U$ is defined by*

$$\mathcal{N}_r(S) \equiv \{S' \in U : \|S - S'\| \leq r\} .$$

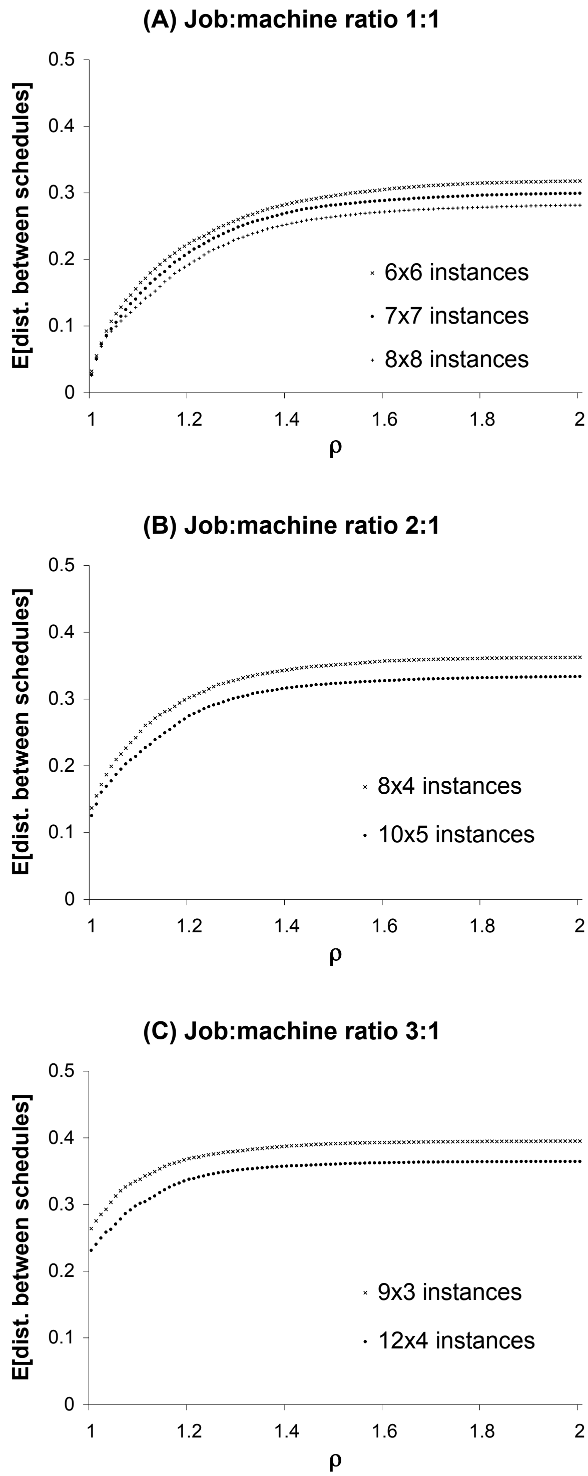

Figure 4: Expected distance between random schedules within a factor ρ of optimality, as a function of ρ. Graphs (A), (B), and (C) depict curves for random instances with $\frac{N}{M}$ = 1, 2, and 3, respectively.

Definition ((r,δ)-valley). *Let I and U be as above, and let r and δ be non-negative integers. A set $V \subseteq U$ is an (r,δ)*-valley *if V has the following two properties.*

1. *For any $S \in V$,*
$$\ell(S) = \min_{\bar{S} \in \mathcal{N}_r(S)} \ell(\bar{S}) \Rightarrow \ell(S) = \min_{\bar{S} \in U} \ell(\bar{S})$$
(i.e., if S that is locally optimal w.r.t. $\mathcal{N}_r$ it must also be globally optimal), while
$$\ell(S) > \min_{\bar{S} \in \mathcal{N}_r(S)} \ell(\bar{S}) \Rightarrow \ell(S) > \min_{\bar{S} \in \mathcal{N}_r(S) \cap V} \ell(\bar{S})$$
(i.e., if S is not locally optimal then some schedule in $\mathcal{N}_r \cap V$ must improve upon it).

2. *For any $\{S_1, S_2\} \subseteq V$ such that $\ell(S_1) = \ell(S_2) = \min_{\bar{S} \in U} \ell(\bar{S})$, $\|S_1 - S_2\| \leq \delta$.*

Definition ((r,δ,p) landscape). *Let I and U be as above, and let S be a random schedule for I. Then I has an (r,δ,p)* landscape *if there exists a $V \subseteq U$ such that*

1. *V is an (r,δ)-valley, and*

2. *$\mathbb{P}[S \in V] \geq p$.*

Any JSSP instance I trivially has an $(M\binom{N}{2}, M\binom{N}{2}, 1)$ landscape. I could be described as having a big valley landscape if I has an (r,δ,p) landscape for small r and δ in combination with p near 1.

In this section we seek to determine the combinations of r and p for which random JSSP instances typically have an $(r, M\binom{N}{2}, p)$ landscape. We do this using a statistic called *neighborhood exactness*, defined below.

Definition ($\mathcal{L}(S,\mathcal{N})$). *Let I, U, and $\mathcal{N}$ be as above, and let S be a schedule for I. The schedule $\mathcal{L}(S,\mathcal{N})$ is obtained by executing the following procedure.*

1. *Let $\mathcal{N}(S) = \{S_0, S_1, \ldots, S_{|\mathcal{N}(S)|}\}$ (where the elements of $\mathcal{N}$ are indexed in a fixed but arbitrary manner).*

2. *Set $i := \min\{j : \ell(S_j) < \ell(S)\}$. If no such i exists, return S; otherwise set $S := S_i$ and go to 1.*

Definition (Neighborhood exactness). *Let I, U, and $\mathcal{N}$ be as above, and let S be a random schedule for I. The exactness of the neighborhood $\mathcal{N}$ on the instance I is the probability that $\mathcal{L}(S,\mathcal{N})$ is a global optimum.*

If the exactness of $\mathcal{N}_r$ is p, then I has an $(r, M\binom{N}{2}, p)$ landscape (let V consist of all schedules S such that $\mathcal{L}(S,\mathcal{N})$ is a global optimum). We will estimate the *expected* exactness of $\mathcal{N}_r$ as a function of r for various combinations of N and M. Examination of the resulting curves will allow us to draw conclusions about how the extent to which typical JSSP landscapes are big valleys changes as a function of $\frac{N}{M}$.

6.2 Estimating Neighborhood Exactness

For given values of N and M, we compute the *expected* exactness of $\mathcal{N}_r$ for $1 \leq r \leq M\binom{N}{2}$ by repeatedly executing the following procedure. To efficiently search for an improving schedule within $\mathcal{N}_r$ we have developed a "radius-limited" branch and bound algorithm.

1. Generate a random N by M JSSP instance I.
2. Using the algorithm of Brucker et al. [5], compute the optimal makespan of I.
3. Let s be a random feasible schedule; let $r = 1$; and let $opt = false$.
4. While $opt = false$ do:
   (a) Starting from s, apply next-descent under the neighborhood N_r to generate a local optimum (each step of next-descent uses our radius-limited branch and bound algorithm). Let s be this local optimum.
   (b) Update the expected exactness of $\mathcal{N}_r$ appropriately, based on whether or not s is a globally optimum.
   (c) If s is a globally optimum, set $opt = true$. Otherwise increment r.
5. For all r' such that $r < r' \leq M\binom{N}{2}$ update the expected exactness of $\mathcal{N}_{r'}$ appropriately.

For each r, the data used to estimate the expected exactness of $\mathcal{N}_r$ are independent, so our estimates are unbiased (data for distinct radii are dependent, however).

Our radius-limited branch and bound algorithm uses the branching rule of Balas [2] combined with the lower bounds and branch ordering heuristic of Brucker et al. [5].

6.3 Results

We use three combinations of N and M with $\frac{N}{M} = \frac{1}{5}$ (3x15, 4x20, and 5x25 instances), three combinations with $\frac{N}{M} = 1$ (6x6, 7x7, and 8x8 instances) and two combinations with $\frac{N}{M} = 5$ (15x3 and 20x4 instances). We generate 1000 random instances for each combination of N and M.

Figure 5 plots expected exactness as a function of neighborhood radius (normalized by the number of disjunctive edges) for each of these three values of $\frac{N}{M}$.

6.4 Discussion

When $\frac{N}{M} = \frac{1}{5}$ or $\frac{N}{M} = 5$, a small (normalized) value of r suffices to ensure that a random local optimum drawn under $\mathcal{N}_r$ is very likely to be a global optimum. Using the methodology of §4, we found that the expected backbone fractions for 3x15, 4x20, and 5x25 instances are 0.94, 0.93, and 0.92, respectively, while the expected distance between global optima was 0.02 in all three cases. This suggests that the typical landscape for $\frac{N}{M} = \frac{1}{5}$ can be described as a big valley. In contrast, the expected backbone fractions for 15x3 and 20x4 instances are near-zero, while the expected

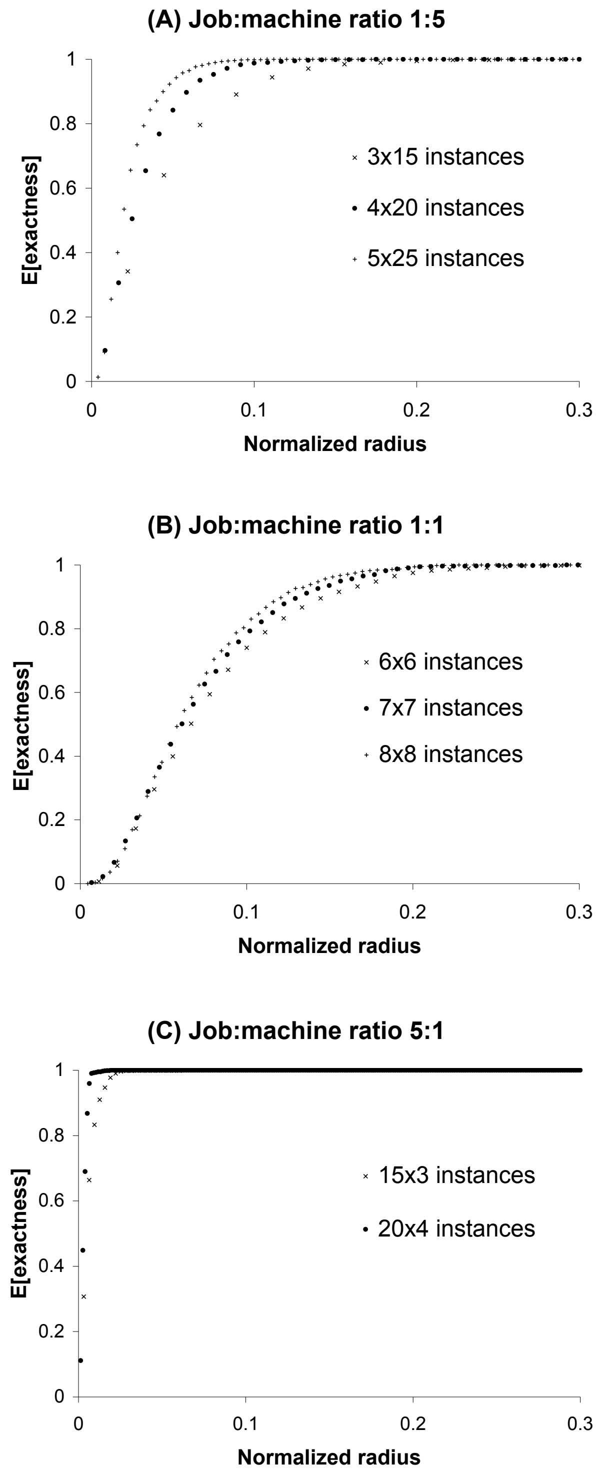

Figure 5: Expected exactness as a function of normalized neighborhood radius. Graphs (A), (B), and (C) depict curves for random instances with $\frac{N}{M} = \frac{1}{5}$, 1, and 5, respectively.

distances between global optima are 0.33 and 0.28, respectively. Thus for $\frac{N}{M} = 5$, the data suggest the existence of many big valleys rather than just one.

For $\frac{N}{M} = 1$, the normalized value of r must be much larger in order to achieve the same expected exactness. The data from §5 show that global optima are fairly tightly clustered when $\frac{N}{M} = 1$, so typical landscapes can still be roughly described as big valleys. However, when $\frac{N}{M} = 1$ the valley is rougher than it is for the more extreme values of $\frac{N}{M}$.

6.5 Analysis

We first establish the behavior of the curves depicted in Figure 5 in the limiting cases $\frac{N}{M} \to 0$ and $\frac{N}{M} \to \infty$. We then use these results to characterize the landscapes of random JSSP instances using the (r, δ, p) notation introduced in §6.1.

The following two lemmas show that as $\frac{N}{M} \to 0$ (resp. $\frac{N}{M} \to \infty$), a random schedule will almost surely be "close" to an optimal schedule. The proofs are given in Appendix A.

Lemma 3. *Let $I_{N,M}$ be a random N by M JSSP instance, and let S be a random schedule for $I_{N,M}$. Let $\hat{S}$ be an optimal schedule for $I_{N,M}$ such that $\|S - \hat{S}\|$ is minimal. Let $f(M)$ be any unbounded, increasing function of M. Then for fixed N, it holds whp (as $M \to \infty$) that $\|S - \hat{S}\| < f(M)$.*

Lemma 4. *Let $I_{N,M}$ be a random N by M JSSP instance, let S be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$ such that $\|S - \hat{S}\|$ is minimal. Then for fixed M and $\epsilon > 0$, it holds whp (as $N \to \infty$) that $\|S - \hat{S}\| < N^{1+\epsilon}$.*

The following are immediate corollaries.

Corollary 3. *For fixed N, the expected exactness of $\mathcal{N}_{f(M)}$ approaches 1 as $M \to \infty$, where $f(M)$ is any unbounded, increasing function of M.*

Corollary 4. *For fixed M and $\epsilon > 0$, the expected exactness of $\mathcal{N}_{N^{1+\epsilon}}$ approaches 1 as $N \to \infty$.*

Because the total number of disjunctive edges is $M\binom{N}{2}$, these two corollaries imply that as $\frac{N}{M} \to 0$ (resp. $\frac{N}{M} \to \infty$), the curve depicted in Figure 5 approaches a horizontal line at a height of 1.

The following two theorems characterize the landscape of random JSSP instances using the (r, δ, p) notation of §6.1.

Theorem 3. *Let $I_{N,M}$ be a random N by M JSSP instance. Let $f(M)$ be any unbounded, increasing function of M. For fixed N and $\epsilon, \epsilon' > 0$, it holds whp (as $M \to \infty$) that $I_{N,M}$ has a (r, δ, p) landscape for $r = f(M)$, $\delta = \epsilon M\binom{N}{2}$ and $p = 1 - \epsilon'$.*

Proof. Let V be the set of all schedules S such that $\mathcal{L}(S, \mathcal{N}_r)$ is a global optimum. It follows by Corollary 3 that whp, $I_{N,M}$ is such that a random schedule S belongs to V with probability at least p. It remains to show that V is an (r, δ)-valley whp. Part 1 of the definition of an (r, δ)-valley is satisfied by the definition of V. Part 2 follows from Theorem 1. □

Theorem 4. *Let $I_{N,M}$ be a random N by M JSSP instance, and let S be a random schedule for $I_{N,M}$. There exists a set $V(I_{N,M}) = \cup_{i=1}^{n} V_i$ of schedules for $I_{N,M}$ such that for fixed M and $\epsilon_1, \epsilon_2, \epsilon_3 > 0$, V has the following properties whp:*

1. $S \in V$;
2. V_i *is an* (r, δ)*-valley with* $r = N^{1+\epsilon_2}$ *and* $\delta = 1 \; \forall i \in [n]$;
3. $n > N!(1 - \epsilon_3)$; *and*
4. $\max_{\{S_1,S_2\} \subseteq V} \|S_1 - S_2\| > \Omega(N^2)$.

Proof. Let $\{\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_n\}$ be the set of globally optimal schedules for $I_{N,M}$, and define $V_i(I_{N,M}) \equiv \{S : \mathcal{L}(S, \mathcal{N}^{1+\epsilon_2}) = \hat{S}_i\}$. Property 1 holds whp by Lemma 4. Property 2 holds by definition of V_i.

The fact that property 3 holds whp is a consequence of Lemma 2. The number of possible choices of the permutation ϕ (used in the statement of Lemma 2) is $N!$. Let f be the number of choices of ϕ that fail to yield a globally optimal schedule. Property 3 can only fail to hold if $f \geq \epsilon_3 N!$. But $\mathbb{E}[f]$ is $o(\exp(-KN)N!)$; hence $f < \epsilon_3 N!$ whp by Markov's inequality.

To establish property 4, choose permutations ϕ_1 and ϕ_2 that list the elements of $[N]$ in reverse order (i.e., $\phi_1(i) = \phi_2(N - i) \; \forall i \in [N]$). These permutations define schedules S_1 and S_2 (via Lemma 2) which are both globally optimal whp. But for any disjunctive edge $e = \{J_1, J_1'\}$ we must have $\vec{e}(S_1) \neq \vec{e}(S_2)$, hence $\|S_1 - S_2\| \geq |\{\{J, J'\} \subseteq I_{N,M} : m(J_1) = m(J_1')\}| \geq \binom{NM^{-1}}{2}$. □

Theorem 3 shows that as $\frac{N}{M} \to 0$, a random JSSP instance almost certainly has an (r, δ, p) landscape where r grows arbitrarily slowly as a function of M, the normalized value of δ (i.e., $\frac{\delta}{M\binom{N}{2}}$) is arbitrarily small, and p is arbitrarily close to 1. In contrast, Theorem 4 shows that as $\frac{N}{M} \to \infty$, a random JSSP instance almost surely does not have an (r, δ, p) landscape unless δ is $\Omega(N^2)$. Instead, the landscape contains $\Omega(N!)$ $(r, 1)$-valleys, with the normalized value of r approaching 0. Random instances with intermediate values of $\frac{N}{M}$ (e.g., $\frac{N}{M} \approxeq 1$) can be seen as an interpolation between these two extremes.

7 Quality of Random Schedules

7.1 Methodology

In this section we examine how the quality of randomly-generated schedules changes as a function of the job:machine ratio. Specifically, for various combinations of N and M, we estimate the expected value of the following four quantities:

(A) the makespan of a random schedule,

(B) the makespan of a locally optimal schedule obtained by starting at a random schedule and applying next-descent using the $\mathcal{N}_1$ move operator,

(C) the makespan of an optimal schedule, and

(D) the lower bound on the makespan of an optimal schedule given by the maximum of the maximum job duration and the maximum machine workload:

$$\max\left(\max_{J \in I_{N,M}} \tau(J), \max_{\bar{m} \in [M]} \sum_{o \in ops(I_{N,M}), m(o)=\bar{m}} \tau(o)\right) .$$

The values of $\frac{N}{M}$ considered in our experiments are those in the set $R = \{\frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1, \frac{3}{2}, 2, 3, 4, 5, 6, 7\}$. We consider all combinations of N and M in the set $S \equiv \bigcup_{r \in R} S_r$, where $S_r \equiv \{(N, M) : \frac{N}{M} = r, \min(N, M) \geq 2, \max(N, M) \geq 6, NM < 1000\}$. For each $(N, M) \in S$, we generate 1000 N by M instances, and compute quantities (A), (B), and (D) for each of these instances. For some combinations $(N, M) \in \bar{S} \subseteq S$, it was also practical to compute quantity (C). We chose $\bar{S}$ so that $|\bar{S} \cap S_r| \geq 4 \; \forall r \in R \setminus \{\frac{3}{2}\}$, while $|\bar{S} \cap S_{\frac{3}{2}}| = 3$.

7.2 Results

Figure 6 plots the mean values of (A), (B), and (C), respectively, against the mean value of (D), for various combinations of N and M. The data points for each combination of N and M are assigned a symbol based on the value of $\frac{N}{M}$. Examining Figure 6, we see that the set of data points for each value of $\frac{N}{M}$ are approximately (though not exactly) collinear. Furthermore, in all three graphs the slope of the line formed by the data points with $\frac{N}{M} = r$ is maximized when $r = 1$, and decreases as r gets further away from 1.

To further investigate this trend, we performed least squares linear regression on the set of data points for each value of $\frac{N}{M}$. The slopes of the resulting lines are shown as a function of $\frac{N}{M}$ in Figure 7.

From examination of Figure 7, it is apparent that

- as the value of $\frac{N}{M}$ becomes more extreme (i.e., approaches either 0 or ∞), the expected makespan of random schedules (resp. random local optima) comes closer to the expected value of the lower bound on makespan; and
- the difference between the expected makespan of random schedules (resp. random local optima) and the expected value of the lower bound on makespan is maximized at a value of $\frac{N}{M} \approxeq 1$.

The first of these two observations suggests that as $\frac{N}{M}$ approaches either 0 or ∞, a random schedule is almost certainly near-optimal. §7.3 contains two theorems that confirm this.

The second of these two observations suggests that the expected difference between the makespan of a random schedule and the makespan of an optimal schedule is maximized at a value of $\frac{N}{M}$ somewhere in the neighborhood of 1. This observation is particularly interesting given the conventional wisdom that square instances of the JSSP (i.e., those with $\frac{N}{M} = 1$) are harder to solve than rectangular ones [7]. We come back to this observation in §7.4.

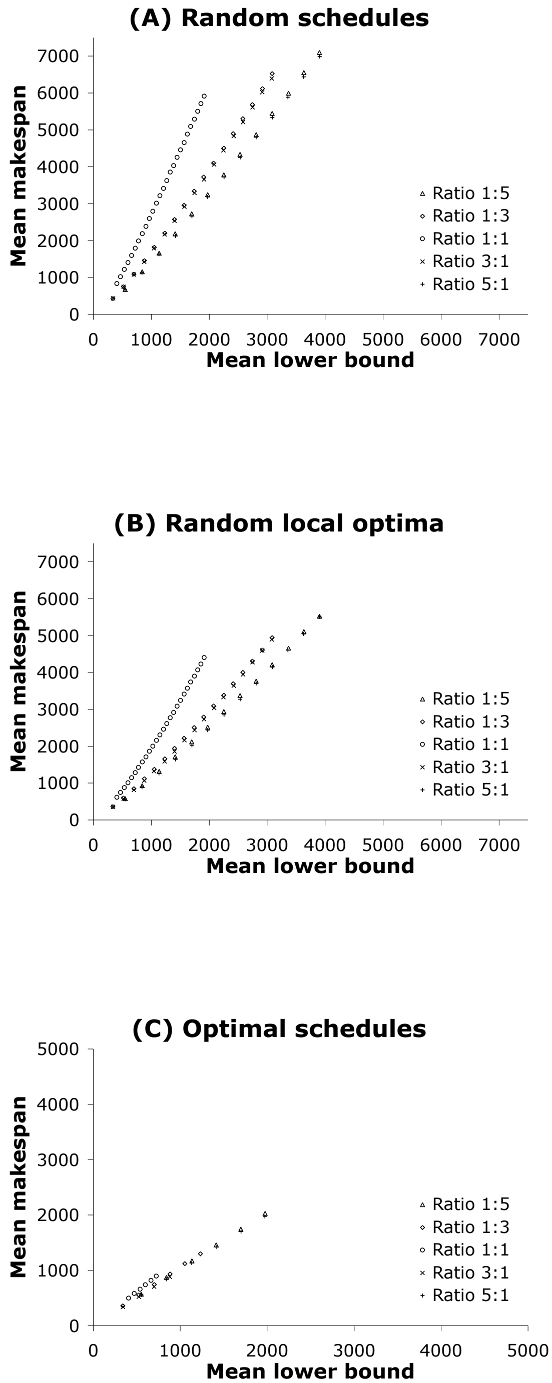

Figure 6: Expected makespan of (A) random schedules, (B) random local optima, and (C) optimal schedules vs. expected lower bound, for various combinations of N and M (grouped by symbol according to $\frac{N}{M}$).

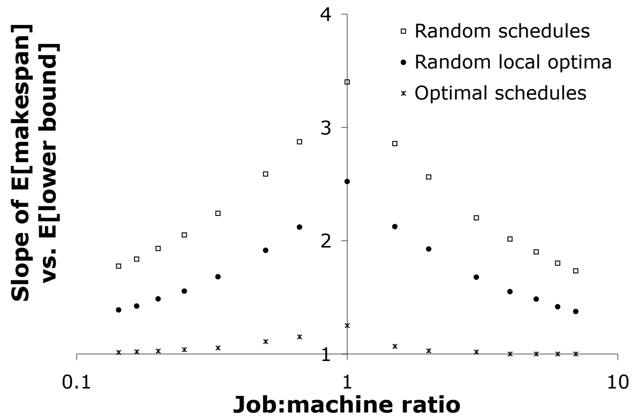

Figure 7: Slope of the least squares fits to the data in Figure 6 (A), (B), and (C) as a function of $\frac{N}{M}$ (includes values of $\frac{N}{M}$ not depicted in Figure 6).

7.3 Analysis

The following two theorems show that, as $\frac{N}{M}$ approaches either 0 or ∞, a random schedule will almost surely be near-optimal (proofs are given in Appendix A).

Theorem 5. *Let $I_{N,M}$ be a random JSSP instance, let S be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$. Then for fixed N and $\epsilon > 0$, it holds whp (as $M \to \infty$) that $\ell(S) \leq (1+\epsilon)\ell(\hat{S})$.*

Theorem 6. *Let $I_{N,M}$ be a random N by M JSSP instance, let S be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$. Then for fixed M and $\epsilon > 0$, it holds whp (as $N \to \infty$) that $\ell(S) \leq (1+\epsilon)\ell(\hat{S})$.*

7.4 Easy-hard-easy pattern of instance difficulty

The proofs of Corollary 1 (resp. Lemma 2) show that as $\frac{N}{M} \to 0$ (resp. $\frac{N}{M} \to \infty$) there exist simple priority rules that almost surely produce an optimal schedule. Moreover, Theorems 5 and 6 show that in these two limiting cases, even a random schedule will almost surely have makespan that is very close to optimal. Thus, both as $\frac{N}{M} \to 0$ and as $\frac{N}{M} \to \infty$, almost all JSSP instances are "easy".

In contrast, for $\frac{N}{M} \approxeq 1$, Figure 7 suggests that random schedules (as well as random local optima) are far from optimal. The literature on the JSSP as well as our own computational experience in using the algorithm of Brucker et al. [5] lead us to believe that random JSSP instances with $\frac{N}{M} \approxeq 1$ are "hard". Thus we conjecture that, as in 3-SAT, typical instance difficulty in the JSSP follows an "easy-hard-easy" pattern as a function of a certain parameter.

8 Conclusions

8.1 Summary of Experimental Results

Empirically, we demonstrated that for low values of the job to machine ratio ($\frac{N}{M}$), low-makespan schedules are clustered in a small region of the search space and the backbone size is high. As $\frac{N}{M}$ increases, low-makespan schedules become dispersed throughout the search space and the the backbone vanishes. For both extremely low and extremely high values of $\frac{N}{M}$, the expected makespan of random schedules comes very close to that of optimal schedules, and the normalized disjunctive graph distance between a random schedule and the nearest optimal schedule becomes very small. The quality of random schedules (resp. random local optima) appears to be the worst at a value of $\frac{N}{M} \approxeq 1$.

§6.4 discussed the implications of our results for the "big valley" picture of JSSP search landscapes. For $\frac{N}{M} \approxeq 1$, we concluded that typical landscapes can be described as a big valley, while for larger values of $\frac{N}{M}$ (e.g., $\frac{N}{M} \geq 3$) there are many big valleys. §7.4 discussed how our data support the idea that JSSP instance difficulty exhibits an "easy-hard-easy" pattern as a function of $\frac{N}{M}$.

8.2 An Overall Picture

Let $I_{N,M}$ be a random N by M JSSP instance; let S_{rand} be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$ such that $\|S_{rand} - \hat{S}\|$ is minimal. Table 1 shows the asymptotic expected values of normalized backbone size, $\frac{\ell(S_{rand})}{\ell(\hat{S})}$, and the normalized distance from S_{rand} to $\hat{S}$ for various values of $\frac{N}{M}$. The values in the first and third columns are provably correct, as shown by theorems 1 through 6. The values in the middle column are conservative conjectures based on our experimental results.

Table 1. Attributes of random JSSP instances as a function of $\frac{N}{M}$.

	$\frac{N}{M} \to 0$	$\frac{N}{M} = k > 0, N \to \infty$	$\frac{N}{M} \to \infty$
$\mathbb{E}\left[\frac{\lvert 1\_backbone(I_{N,M})\rvert}{M\binom{N}{2}}\right]$	1	$\in [0,1]$	0
$\mathbb{E}\left[\frac{\ell(S_{rand})}{\ell(\hat{S})}\right]$	1	> 1	1
$\mathbb{E}\left[\frac{\|S_{rand}-\hat{S}\|}{M\binom{N}{2}}\right]$	0	> 0	0

References

[1] Dimitris Achlioptas and Yuval Peres. The threshold for random k-sat is $2^k \log 2 - O(k)$. *Journal of the AMS*, 17:947–973, 2004.

[2] Egon Balas. Machine sequencing via disjunctive graphs: An implicit enumeration algorithm. *Operations Research*, 17:1–10, 1969.

[3] J.E. Beasley. OR-library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41(11):1069–1072, 1990.

[4] Kenneth D. Boese, Andrew B. Kahng, and Sudhukar Muddu. A new adaptive multi-start technique for combinatorial global optimizations. *Operations Research Letters*, 16:101–113, 1994.

[5] Peter Brucker, Bernd Jurisch, and Bernd Sievers. A branch and bound algorithm for the job-shop scheduling problem. *Discrete Applied Mathematics*, 49(1-3):107–127, 1994.

[6] Peter Cheeseman, Bob Kanefsky, and William M. Taylor. Where the really hard problems are. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, IJCAI-91, Sidney, Australia*, pages 331–337, 1991.

[7] H. Fisher and G. L. Thompson. Probabilistic learning combinations of local job-shop scheduling rules. In J. F. Muth and G. L. Thompson, editors, *Industrial Scheduling*, pages 225–251. Prentice-Hall, Englewood Cliffs, NJ, 1963.

[8] Fred Glover and Manuel Laguna. *Tabu Search*. Kluwer Academic Publishers, Boston, MA, 1997.

[9] A.S. Jain and S. Meeran. A state-of-the-art review of job-shop scheduling techniques. Technical report, Department of Applied Physics, Electronic and Mechanical Engineering, University of Dundee, Dundee, Scotland, 1998.

[10] Albert Jones and Luis C. Rabelo. Survey of job shop scheduling techniques. Technical report, National Institute of Standards and Technology, Gaithersburg, MD, 1998.

[11] Yong-Hyuk Kim and Byung-Ro Moon. Investigation of the fitness landscapes in graph bipartitioning: An empirical study. *Journal of Heuristics*, 10:111–133, 2004.

[12] H.R. Lourenço, O. Martin, and T. Stützle. Iterated local search. In F. Glover and G. Kochenberger, editors, *Handbook of Metaheuristics*. Kluwer Academic Publishers, Boston, MA, 2003.

[13] Dorothy L. Mammen and Tad Hogg. A new look at the easy-hard-easy pattern of combinatorial search difficulty. *Journal of Artificial Intelligence Research*, 7:47–66, July 1997.

[14] Olivier C. Martin, Rémi Monasson, and Ricardo Zecchina. Statistical mechanics methods and phase transitions in combinatorial problems. *Theoretical Computer Science*, 265(1-2):3–67, 2001.

[15] Dirk C. Mattfeld. *Evolutionary Search and the Job Shop: Investigations on Genetic Algorithms for Production Scheduling*. Physica-Verlag, Heidelberg, 1996.

[16] M. Mézard and G. Parisi. A replica analysis of the traveling salesman problem. *Journal de Physique*, 47:1285–1296, 1986.

[17] Rémi Monasson, Riccardo Zecchina, Scott Kirkpatrick, Bart Selman, and Lidror Troyansky. Determining computational complexity from characteristic 'phase transitions'. *Nature*, 400:133–137, 1999.

[18] Eugeniusz Nowicki and Czeslaw Smutnicki. A fast taboo search algorithm for the job-shop problem. *Management Science*, 42(6):797–813, 1996.

[19] Eugeniusz Nowicki and Czeslaw Smutnicki. Some new ideas in TS for job shop scheduling. Technical Report 50/2001, University of Wroclaw, 2001.

[20] Eugeniusz Nowicki and Czeslaw Smutnicki. An advanced tabu search algorithm for the job shop problem. *Journal of Scheduling*, 8:145–159, 2005.

[21] C. R. Reeves and T. Yamada. Genetic algorithms, path relinking, and the flowshop sequencing problem. *Evolutionary Computation*, 6:45–60, 1998.

[22] B. Roy and B. Sussmann. Les problèmes dordonnancement avec contraintes disjonctives. Note D.S. no. 9 bis, SEMA, Paris, France, Décembre, 1964.

[23] John Slaney and Toby Walsh. Backbones in optimization and approximation. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI-2001)*, pages 254–259, 2001.

[24] Joel Spencer. Modern probabilistic methods in combinatorics. `http://www.cs.nyu.edu/cs/faculty/spencer/papers/stirlingtalk.pdf`, 2005.

[25] Matthew J. Streeter and Stephen F. Smith. Supplemental material for ICAPS 2005 paper 'Characterizing the distribution of low-makespan schedules in the job shop scheduling problem'. `http://www.cs.cmu.edu/~matts/icaps_2005`, 2005.

[26] É. Taillard. Parallel taboo search techniques for the job shop scheduling problem. *ORSA Journal on Computing*, 6:108–117, 1994.

[27] P.J.M. van Laarhoven, E.H.L. Aarts, and J.K. Lenstra. Job shop scheduling by simulated annealing. *Operations Research*, 40(1):113–125, 1992.

[28] Jean-Paul Watson, Laura Barbulescu, L. Darrell Whitley, and Adele Howe. Contrasting structured and random permutation flow-shop scheduling problems: search-space topology and algorithm performance. *INFORMS Journal on Computing*, 14(2):98–123, 2002.

[29] Jean-Paul Watson, J. Christopher Beck, Adele E. Howe, and L. Darrell Whitley. Toward an understanding of local search cost in job-shop scheduling. In A. Cesta, editor, *Proceedings of the Sixth European Conference on Planning*, 2001.

Appendix A: Additional Proofs

We make use of the following inequality [24].

Azuma's Perimetric Inequality (A.P.I.). *Let $X = (X_1, X_2, \dots, X_n)$ be a vector of n independent random variables. Let the function $f(x)$ take as input a vector $x = (x_1, x_2, \dots, x_n)$, where x_i is a realization of X_i for $i \in [n]$, and produce as output a real number. Suppose that for some $\beta > 0$ it holds that for any two vectors x and x' that differ on at most one coordinate,*

$$|f(x) - f(x')| \leq \beta \, .$$

Then for any $\alpha > 0$,

$$\mathbb{P}[X > \mathbb{E}[X] + \alpha\sqrt{n}] \leq \exp\left(-\frac{\alpha^2}{2\beta^2}\right) \, .$$

The same inequality holds for $\mathbb{P}[X \leq \mathbb{E}[X] - \alpha\sqrt{n}]$.

Lemma 2. *Let $I_{N,M}$ be a random N by M JSSP instance. Let ϕ be an arbitrary permutation of $[N]$, and consider the priority rule π defined by*

$$\pi(I_{N,M}) = T_1 T_2 \dots T_M$$

where

$$T_i = (J_i^{\phi(1)}, J_i^{\phi(2)}, \dots, J_i^{\phi(N)}) \, .$$

Then for fixed M, there exists a $K > 0$ such that with probability $1 - o(\exp(-KN))$, the schedule $\hat{S} = S(\pi, I_{N,M})$ has the property that

$$\hat{S}(o) = \hat{S}^+(\mathcal{M}(o)) \; \forall o \in ops(I_{N,M}) \, .$$

Proof. Let $I_{N,M} = \{J^1, J^2, \dots, J^N\}$, and assume without loss of generality that $\phi(i) = i \; \forall i \in [N]$. Let the (not necessarily feasible) schedule $\bar{S}$ be defined by

$$\bar{S}(J_i^k) = \tau\left(\{J_h^j \in ops(I_{N,M}) : (h < i \vee j < k) \wedge m(J_h^j) = m(J_i^k)\}\right)$$

for $J_i^k \in ops(I_{N,M})$.

$\bar{S}$ clearly has the property described in the statement of the lemma. It is straightforward to check that if $\bar{S}$ is feasible, it is identical to $\hat{S}$. Thus it remains to show that for some $K > 0$, $\bar{S}$ is infeasible with probability $o(\exp(-KN))$. By definition, $\bar{S}$ will be infeasible iff. there is some operation $o \in ops(I_{N,M})$ such that $\bar{S}^+(\mathcal{M}(o)) < \bar{S}^+(\mathcal{J}(o))$. It suffices to show that for any $o \in ops(I_{N,M})$, $\mathbb{P}[\bar{S}^+(\mathcal{M}(o)) < \bar{S}^+(\mathcal{J}(o))] \leq \exp(-\epsilon N)$, for some fixed $\epsilon > 0$.

If $i = 1$, $\mathbb{P}[\bar{S}^+(\mathcal{M}(J_i^k)) < \bar{S}^+(\mathcal{J}(J_i^k))] = 0$. Otherwise,

$$\mathbb{E}[\bar{S}^+(\mathcal{M}(J_i^k)) - \bar{S}^+(\mathcal{J}(J_i^k))] = \mu\frac{(i-1)N+k-1}{M} - \mu\frac{(i-2)N+k-1}{M} = \mu\frac{N}{M} \ .$$

The value of $\bar{S}^+(\mathcal{M}(J_i^k)) - \bar{S}^+(\mathcal{J}(J_i^k))$ is a function of N independent events (namely, the definition of each of the N jobs in $I_{N,M}$), and altering a particular job changes the value of this expression by at most τ_{max}. Thus by A.P.I.,

$$\mathbb{P}[\bar{S}^+(\mathcal{M}(J_i^k)) < \bar{S}^+(\mathcal{J}(J_i^k))] \leq \exp\left(-\frac{\mu^2 N}{2(M\tau_{max})^2}\right) \ .$$

$\square$

Lemma 3. *Let $I_{N,M}$ be a random N by M JSSP instance, and let S be a random schedule for $I_{N,M}$. Let $\hat{S}$ be an optimal schedule for $I_{N,M}$ such that $\|S - \hat{S}\|$ is minimal. Let $f(M)$ be any unbounded, increasing function of M. Then for fixed N, it holds whp (as $M \to \infty$) that $\|S - \hat{S}\| < f(M)$.*

Proof. Let $\hat{J} = \arg\max_{J \in I_{N,M}} \tau(J)$. Let $\bar{S} \equiv \pi_{max}(I_{N,M})$ be the schedule obtained by the priority rule π_{max} (discussed in the proof of Corollary 1) that first schedules the operations in $\hat{J}$, then schedules the remaining operations of $I_{N,M}$ in some arbitrary order. The proof of Corollary 1 showed that for any J, $\mathbb{E}[\Delta_J^{\bar{S}}]$ is $O(N^2)$. Thus it holds whp that $\Delta_J^{\bar{S}} < \log(f(M)) \ \forall J$. The procedure used to produce S is a mixture of instance-independent priority rules, each subject to Lemma 1. Thus for any J, $\mathbb{E}[\Delta_J^S]$ is $O(N)$, so whp $\Delta_J^S < \log(f(M)) \ \forall J$.

Let $O_{near}(J_i) = \{J'_j : J' \neq J, |\sum_{i'<i} \tau(J_{i'}) - \sum_{j'<j} \tau(J'_{j'})| < 2\log(f(M))\}$. ($O_{near}(J_i)$ is the set of operations that would be scheduled "near" in time to J_i if resource constraints were ignored.) Let $E_{near} = \{e = \{J_i, J'_j\} \in E(I_{N,M}) : J'_j \in O_{near}(J_i)\}$. Under the assumptions of the previous paragraph (each of which hold whp), $\|S - \bar{S}\| \leq |E_{near}|$. For any J_i, $\mathbb{E}[|O_{near}(J_i)|]$ is $O(N \log f(M))$, and each $J'_j \in O_{near}(J_i)$ has probability $\frac{1}{M}$ of using the same machine as J_i. It follows that $\mathbb{E}[|E_{near}|]$ is $O(N^2 \log f(M))$. Thus $|E_{near}|$ does not exceed $f(M)$ whp.

$\square$

For the purpose of the remaining proofs, it is convenient to introduce some additional notation. Let $T = (T_1, T_2, \ldots, T_{|T|})$ be a sequence of operations. We define

- $T_{(i_1,i_2]} \equiv \{T_i : i_1 < i \leq i_2\}$, and
- $T_{(i_1,i_2]}^{\bar{m}} \equiv \{T_i \in T_{(i_1,i_2]} : m(T_i) = \bar{m}\}$.

Lemma 4. *Let $I_{N,M}$ be a random N by M JSSP instance, let S be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$ such that $\|S - \hat{S}\|$ is minimal. Then for fixed M and $\epsilon > 0$, it holds whp (as $N \to \infty$) that $\|S - \hat{S}\| < N^{1+\epsilon}$.*

Proof. Let $T = \pi_{rand}(I_{N,M})$ be the sequence used to construct S, and let T_i denote the i^{th} operation in T. Consider the schedule $\bar{S}$ defined by the following procedure:

1. Set $\bar{S}(o) := \infty \ \forall o \in ops(I_{N,M})$.
2. Set $Q := ()$. Let Q_j denote the j^{th} operation in Q.
3. Let the function $ready(o)$ return true if $\bar{S}^+(\mathcal{M}(o)) \geq \bar{S}^+(\mathcal{J}(o))$, false otherwise.
4. For i from 1 to NM do:
   (a) If $ready(T_i)$, set $\bar{S}(o) := \bar{S}^+(\mathcal{M}(T_i))$. Otherwise append T_i onto Q.
   (b) For j from 1 to $|Q|$ do:
      i. If $ready(Q_j)$, set $\bar{S}(Q_j) := \bar{S}^+(\mathcal{M}(Q_j))$ and remove Q_j from Q.
5. Schedule any remaining operations of Q in a manner to be specified.

The construction of $\bar{S}$ is just like the construction of S, except for the manipulations involving Q. The purpose of Q is to delay the scheduling of any operation o that, if scheduled immediately, might produce a schedule in which $\bar{S}(o) > \bar{S}^+(\mathcal{M}(o))$. We first show that $\|S - \bar{S}\| < N^{1+\epsilon}$ whp; then we show that $\bar{S}$ is optimal whp.

Let Q^i denote Q as it exists after i iterations of step 4 have been performed. Let $q(o) = \sum_{i=1}^{NM} |o \cap Q^i|$ be the number of iterations during which $o \in Q$. We claim that $\|S - \bar{S}\| \leq \sum_{o \in ops(I_{N,M})} q(o) + (N-1)|Q^{NM}|$. Letting $E^{\neq} = \{e \in E(I_{N,M}) : \vec{e}(\bar{S}) \neq \vec{e}(S)\}$, we have

$$\begin{aligned} \|S - \bar{S}\| &= |\{e \in E^{\neq} : e \cap Q^{NM} = \emptyset\}| + |\{e \in E^{\neq} : e \cap Q^{NM} \neq \emptyset\}| \\ &\leq |\{e \in E^{\neq} : e \cap Q^{NM} = \emptyset\}| + (N-1)|Q^{NM}| \end{aligned}$$

so it suffices to show $|\{e \in E^{\neq} : e \cap Q^{NM} = \emptyset\}| \leq \sum_{o \in ops(I_{N,M})} q(o)$. To see this, let $e = \{o_1, o_2\} \in E^{\neq}$ be such that $e \cap Q^{NM} = \emptyset$. We must have $q(o_1) + q(o_2) > 0$. We charge e to the operation in $\{o_1, o_2\}$ that was inserted into Q first. It is easy to see that an operation can be charged for at most one edge per iteration it spends in Q, establishing our claim. Thus it suffices to show that $\|S - \bar{S}\| \leq \sum_{o \in ops(I_{N,M})} q(o) + (N-1)|Q^{NM}| \leq N^{1+\epsilon}$ whp.

We divide the construction of S into $n = MN^{\frac{1}{2}-\epsilon'}$ epochs, each consisting of $N^{\frac{1}{2}+\epsilon'}$ iterations of step 4, for a to-be-specified $\epsilon' > 0$. Let z_j denote the number of iterations of step 4 that occur before the end of the j^{th} epoch, with $z_j = 0$ for $j \leq 0$ by convention. Let

- $C_j^{\bar{m}} \equiv T_{(0,z_j]}^{\bar{m}} \setminus Q^{z_j}$ be the set of operations that have been scheduled to run on $\bar{m}$ by the end of the j^{th} epoch; and

- $O_{near} \equiv \bigcup_{j\in[n]}\{o \in T_{(z_{j-1},z_j]} : \mathcal{J}(o) \in T_{(z_{j-(M+2)},z_j]}\}$ be the set of operations whose job-predecessor belongs to a nearby epoch.

For any $i \in [NM]$, $\mathbb{P}[T_i \in O_{near}] \leq (M+2)N^{-\frac{1}{2}+\epsilon'}$. Thus for any $j \in [n]$, $\mathbb{E}[|O_{near} \cap T_{(z_{j-1},z_j]}|] \leq (M+2)N^{2\epsilon'}$. Using A.P.I. it is straightforward to show that whp,

$$|O_{near} \cap T_{(z_{j-1},z_j]}| \leq N^{\frac{1+\epsilon'}{2}} \ \forall j \in [n] \,. \tag{8.1}$$

We claim that whp, the following statements hold $\forall j \in [n]$:

$$\bigcup_{i \leq z_j} Q^i \subseteq O_{near} \,, \tag{8.2}$$

$$J \cap Q^{z_{j-1}} \neq \emptyset \Rightarrow |J \cap Q^{z_{j-1}} \cap Q^{z_j}| < |J \cap Q^{z_{j-1}}| \qquad \forall J \in I_{N,M} \,, \tag{8.3}$$

$$Q^{z_j} \cap Q^{z_{j-M}} = \emptyset \text{ , and} \tag{8.4}$$

$$|Q^{z_j}| \leq MN^{\frac{1+\epsilon'}{2}} \,. \tag{8.5}$$

We prove this by induction, where each step of the induction fails with exponentially small probability. For $j = 0$, (8.3) and (8.4) hold trivially. (8.2) is true because the operations in $T_{(0,z_1]} \setminus O_{near}$ are the first operations in their jobs, hence cannot be added to Q. (8.5) then follows from (8.2) and (8.1).

Consider the case $j > 0$. To show (8.2), let o be an arbitrary operation in $T_{(z_{j-1},z_j]} \setminus O_{near}$. By the induction hypothesis (specifically, equation (8.4)), $\mathcal{J}(o) \in C_{j-2}^{m(\mathcal{J}(o))}$. Thus $q(o) > 0 \Rightarrow \tau\left(C_{j-2}^{m(\mathcal{J}(o))}\right) > \tau\left(C_{j-1}^{m(o)}\right)$. By the induction hypothesis,

$$\tau\left(C_{j-1}^{m(o)}\right) - \tau\left(C_{j-2}^{m(\mathcal{J}(o))}\right) \geq \tau\left(T_{(0,z_{j-1}]}^{m(o)}\right) - MN^{\frac{1+\epsilon'}{2}} - \tau\left(T_{(0,z_{j-2}]}^{m(\mathcal{J}(o))}\right) \,.$$

Letting Δ denote the right hand side of this inequality, we have $\mathbb{E}[\Delta] = \frac{1}{M}N^{\frac{1}{2}+\epsilon'} - MN^{\frac{1+\epsilon'}{2}}$, and A.P.I. can be used to show that for some $K > 0$ independent of N, $\mathbb{P}[\Delta < 0] \leq \exp(-\frac{1}{K}N^{\epsilon'})$. Thus (8.2) holds with probability at least $1 - \exp(-\frac{1}{K}N^{\epsilon'})$.

To show (8.3), let J be such that $J \cap Q^{z_{j-1}} \neq \emptyset$, and let $J_i \in Q^{z_{j-1}}$ be chosen so that i is minimal. Then $\mathcal{J}(J_i) \in C_{j-1}^{m(\mathcal{J}(J_i))}$. Thus $J_i \in Q_{z_j} \Rightarrow \tau\left(C_{j-1}^{m(\mathcal{J}(J_i))}\right) > \tau\left(C_j^{m(J_i)}\right)$. By (8.1), (8.2), and the induction hypothesis (equation (8.5)), $|Q^{z_j}| \leq (M+1)N^{\frac{1+\epsilon'}{2}}$. Using the same technique as above, we can show that (8.3) holds with probability at least $1 - \exp(-\frac{1}{K}N^{\epsilon'})$ for some $K > 0$ independent of N.

(8.3) implies (8.4). (8.2) and (8.4) together with (8.1) imply (8.5). Thus whp, (8.2) through (8.5) hold $\forall j \in [n]$.

By (8.2) and (8.4), we have

$$\mathbb{E}\left[\sum_{o \in ops(I_{N,M})} q(o)\right] \leq \mathbb{E}[|O_{near}|]MN^{\frac{1}{2}+\epsilon'} \leq M^2(M+2)N^{1+2\epsilon'}$$

and also

$$\mathbb{E}[|Q^{NM}|] \leq \mathbb{E}[|T_{(z_{n-M}, z_n]} \cap O_{near}|] \leq (M+2)N^{2\epsilon'}$$

so setting $\epsilon' = \frac{\epsilon}{3}$ gives $\|S - \bar{S}\| \leq \sum_{o \in ops(I_{N,M})} q(o) + (N-1)|Q^{NM}| \leq N^{1+\epsilon}$ whp.

It remains to show that $\bar{S}$ is optimal whp. The operations scheduled prior to step 5 do not cause any idle time on any machine, so it is only the operations in Q^{NM} that can cause $\bar{S}$ to be sub-optimal. Let $\tau(\bar{m}) \equiv \tau(\{o \in ops(I_{N,M}) : m(o) = \bar{m}\})$ denote the workload of machine $\bar{m}$. Let $\hat{m} = \arg\max_{\bar{m} \in [M]} \tau(\bar{m})$. Then the following hold whp.

- The set $Z^{\hat{m}} \equiv T^{\hat{m}}_{(NM - 2MN^{\frac{1}{4}}, NM]}$ consists of operations belonging to jobs that use $\hat{m}$ last.
- $\mu N^{\frac{1}{4}} \leq \tau(Z^{\hat{m}})$ and $\tau(Z^{\hat{m}}) \leq \tau(\hat{m}) - \tau(\bar{m})\ \forall \bar{m} \neq \hat{m}$.

Thus whp it holds that prior to the execution of step 5, $\bar{S}$ contains a period of length at least $\tau(Z^{\hat{m}}) \geq \mu N^{\frac{1}{4}}$ during which the only operations being processed are those in $Z^{\hat{m}}$, where $\{o \in ops(I_{N,M}) : \mathcal{J}(o) \in Z^{\hat{m}}\} = \emptyset$. Assuming $|Q^{NM}| < N^{3\epsilon'}$ (holds whp), we can always schedule the operations in Q^{NM} so as to guarantee $\ell(\bar{S}) = \tau(\hat{m})$, which implies $\bar{S}$ is optimal.

□

Theorem 5. *Let $I_{N,M}$ be a random JSSP instance, let S be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$. Then for fixed N and $\epsilon > 0$, it holds whp (as $M \to \infty$) that $\ell(S) \leq (1+\epsilon)\ell(\hat{S})$.*

Proof. The procedure used to construct random schedules is a mixture of instance-independent priority rules, each subject to Lemma 1. Thus for each J, $\mathbb{E}[\Delta^S_J]$ is $O(N)$. Thus $\ell(S) - \ell(\hat{S}) \leq \sum_J \Delta^S_J = O(N^2)$, and thus does not exceed $\epsilon\ell(\hat{S})$ whp. □

Theorem 6. *Let $I_{N,M}$ be a random N by M JSSP instance, let S be a random schedule for $I_{N,M}$, and let $\hat{S}$ be an optimal schedule for $I_{N,M}$. Then for fixed M and $\epsilon > 0$, it holds whp (as $N \to \infty$) that $\ell(S) \leq (1+\epsilon)\ell(\hat{S})$.*

Proof. Let $T = \pi_{rand}(I_{N,M})$ be the sequence used to construct S, and let T_i be the i^{th} operation in T. Rather than analyze S directly, we analyze a schedule $\bar{S}$ defined by the following procedure:

1. Set $t := 0$.
2. For i from 1 to NM do:
   (a) Set $\bar{S}(T_i) = \max(t, \bar{S}^+(\mathcal{J}(T_i)), \bar{S}^+(\mathcal{M}(T_i)))$.
   (b) If $\bar{S}^+(\mathcal{J}(T_i)) > \bar{S}^+(\mathcal{M}(T_i))$, set $t = \max_{i' \leq i} \bar{S}^+(T_{i'})$.

The procedure is identical to the one used to construct S, except that, whenever an operation T_i is assigned a start time $\bar{S}(T_i) > \bar{S}^+(\mathcal{M}(T_i))$, the procedure inserts artificial delays into the schedule in order to re-synchronize the machines. For any T, it is clear that $\ell(S) \leq \ell(\bar{S})$. Thus, it suffices to show that $\ell(\bar{S}) \leq (1+\epsilon)\ell(\hat{S})$ whp.

We divide the construction of $\bar{S}$ into n epochs, where each update to t (in step 2(b)) defines the beginning of a new epoch. Let z_i be the number of operations scheduled before the end of the i^{th} epoch, with $z_0 = 0$ by convention. Let $t_i = \max_{i' \leq z_i} S^+(o_{i'})$ be the (updated) value of t at the end of the i^{th} epoch. Define $\Delta_i \equiv \sum_{\bar{m}=1}^{M} t_i - \max_{i'<i, m(T_{i'})=\bar{m}} S^+(T_{i'})$. Then $\ell(\bar{S}) - \ell(\hat{S}) \leq \sum_{i=1}^{n} \Delta_i$, so It suffices to show that $\sum_{i=1}^{n} \Delta_i \leq \epsilon \ell(\hat{S})$ whp.

Let $I = [n]$, and let $L = \{i \in I : z_i - z_{i-1} \geq N^{\frac{2}{7}}\}$. We first consider $\sum_{i \in L} \Delta_i$; then we consider $\sum_{i \in I \setminus L} \Delta_i$.

Let i_1 and i_2 be arbitrary integers with $0 \leq i_1, i_2 \leq NM$ and $i_2 - i_1 \geq N^{\frac{2}{7}}$. Let $\bar{\tau} = \tau(T^{\bar{m}}_{(i_1,i_2]})$. Then $\mathbb{E}[\bar{\tau}] = \mu \frac{i_2 - i_1}{M}$. For any T, $\bar{\tau}$ is a function of the outcome of at most $i_2 - i_1$ events (namely, the definition of each of the jobs in $\{J : J \cap T_{(i_1,i_2]} \neq \emptyset\}$), each of which alters the value of $\bar{\tau}$ by at most τ_{max}. It follows by A.P.I. that

$$\mathbb{P}[|\bar{\tau} - \mathbb{E}[\bar{\tau}]| > N^{\epsilon'}\sqrt{i_2 - i_1}] \leq 2\exp\left(-\frac{N^{2\epsilon'}}{2\tau_{max}^2}\right)$$

for any $\epsilon' > 0$. Thus, it holds whp that $|\bar{\tau} - \mathbb{E}[\bar{\tau}]| \leq N^{\epsilon}\sqrt{i_2 - i_1}$ for all possible choices of i_1 and i_2. In particular, whp we have $\Delta_i \leq 2MN^{\epsilon'}\sqrt{z_i - z_{i-1}}\ \forall i$, which implies $\sum_{i \in K} \Delta_i \leq 2MN^{\frac{6}{7}+\epsilon'}$.

Now consider $\sum_{i \in I \setminus L} \Delta_i$. It can be easily shown that the probability that an arbitrary set of at most $N^{\frac{2}{7}}$ consecutive operations in T contains two operations from the same job is at most $N^{-\frac{3}{7}}$, so $\mathbb{E}[|I \setminus L|] \leq N^{\frac{4}{7}}$. Clearly $\Delta_i \leq \tau_{max} N^{\frac{2}{7}}\ \forall i \in I \setminus L$, so $\mathbb{E}[\sum_{i \in I \setminus L} \Delta_i]$ is $O(N^{\frac{6}{7}})$.

Thus $\mathbb{E}[\sum_{i \in I} \Delta_i]$ is $O(N^{\frac{6}{7}+\epsilon'})$ for any $\epsilon' > 0$, so $\sum_{i \in I} \Delta_i \leq N^{\frac{6}{7}+2\epsilon'}$ whp, while it is easy to see that $\ell(\hat{S}) \geq \mu\frac{N}{2}$ whp.

$\square$