Additional Results and Extensions for the paper "Probabilistic bounds on the k-Traveling Salesman Problem and the Traveling Repairman Problem"

M. Blanchard[1], A. Jacquillat[1], and P. Jaillet[1]

[1]*Massachusetts Institute of Technology, Cambridge, MA, USA*

We study two variants of the classical traveling salesman problem (TSP). Given n points, the TSP seeks a tour of minimal length visiting all n points. In contrast, we focus on

- the k-TSP which seeks a path of minimal length visiting k out of n points, where $k \leq n$. Formally, if $x_1, \ldots, x_k$ is the service order, the objective to minimize is the path length

$$\sum_{i=1}^{k-1} |x_{i+1} - x_i|.$$

- the traveling repairman problem (TRP) which seeks a tour visiting all n points that minimize the sum of latencies (or waiting time) for each point. Formally, if $x_1, \ldots, x_n$ defines a service order, the latency at point x_i is defined as $l_i = \sum_{j=1}^{i-1} |x_{j+1} - x_j|$ and the objective is to minimize the total latency

$$\sum_{i=1}^{n} l_i = \sum_{i=1}^{n-1} (n-i)|x_{i+1} - x_i|.$$

We consider a probabilistic setting where n points $X_1, \ldots, X_n$ are sampled independently and identically from some distribution on a compact $\mathcal{K} \subset \mathbb{R}^2$.

In [2], we provided *constant-factor* probabilistic approximations of both problems, i.e., bounds on the expected optimal objective value that hold within a universal constant factor, as well as constant-factor approximation algorithms. Precisely, we show that the optimal length of the k-TSP path (non-asymptotically) grows at a rate of $\Theta\left(k/n^{\frac{1}{2}(1+\frac{1}{k-1})}\right)$ and that a constant-factor approximation scheme can be obtained by solving the TSP in a high-concentration zone, leveraging large deviations of local point concentration. Next, we show that the optimal TRP objective follows an asymptotic rate $\Theta(n\sqrt{n})$ with a prefactor that depends on the density f of the absolutely-continuous part of the point distribution. This generalizes the classical Beardwood-Halton-Hammersley theorem to the latency-minimization objective in the TRP. The resulting constant-factor approximation scheme visits local regions of the space by decreasing order of probability density f. Last, we propose fairness-enhanced versions of the k-TSP and the TRP to balance efficiency and fairness.

In this companion report, we provide two additional contributions.

1. We extend the k-TSP results to the case with general densities. In Section 1, we show that the results obtained in [2] with continuous densities can be extended via smoothing techniques. We also discuss the case of $k = \Omega(n)$, in which case the k-TSP path becomes non-local and recovers similar behavior to that of the TRP tour—visiting zones by decreasing order of density until k points are visited.

2. For the TRP, we propose a utility-based notion of fairness in Section 2. Instead of assuming that the dissatisfaction (or negative utility) of customers is linear in their latency/waiting time, we consider the case where the utility is a convex function Ψ of their latency. A fair solution aims to minimize total dissatisfaction, which we refer to as the Ψ-TRP solution. For polynomial functions Ψ, we give constant-factor approximations of the optimal Ψ-TRP objective, thus extending the TRP bounds to non-linear utility. Further, we show that the approximation scheme for the TRP given in [2] can be efficiently adapted to obtain constant-factor approximations in the Ψ-TRP.

1 Generalisations of probabilistic bounds for the k-TSP

In the main paper, we provide probabilistic bounds for the k-TSP when points are sampled independently from a distribution with continuous density on a compact. In this section, we present a natural extension of this result to distributions with general densities f on a compact. In particular, the density f is allowed to diverge on a zero-measure set. To this end, we use the notion of Lebesgue derivative $\tilde{f}$, defined as the local average value of f on centered balls. Intuitively, $\tilde{f}$ is a smoothed version of the density f. For instance, if f is continuous then $\tilde{f} = f$. Formally, the Lebesgue derivative is defined as follows:

$$\tilde{f}(x) := \lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f, \quad \forall x,$$

where $|B(x,r)|$ denotes the volume of a centered ball at x of radius r. The Lebesgue differentiation theorem states that this limit exists and that $\tilde{f}$ and f coincide almost everywhere. By construction, the maximum density of points sampled according to f cannot exceed $\|\tilde{f}\|_\infty$. Because f and $\tilde{f}$ coincide almost everywhere, if $\|\tilde{f}\|_\infty < \infty$, the same proof as for continuous densities gives this non-asymptotic lower bound for the length of the k-TSP, where f has simply been replaced by $\tilde{f}$.

Proposition 1.1. *Assume n vertices are drawn independently, on a compact space $\mathcal{K}$, according to a density f such that its Lebesgue derivative $\tilde{f}$ is bounded on $\mathcal{K}$. Denote by $l_{TSP}(k,n)$ the length of the $k-TSP$ on these n vertices, where $2 \le k \le n$. There exists a universal constant $c > 0$ such that*

$$\mathbb{E}[l_{TSP}(k,n)] \ge c \frac{k-1}{(\|\tilde{f}\|_\infty n)^{\frac{1}{2}(1+\frac{1}{k-1})}} \mathcal{A}_{\mathcal{K}}^{-\frac{1}{2(k-1)}}.$$

For the upper bound, we provide similar asymptotic results, which match the lower bound whenever $k \to \infty$ and $k = o(n)$.

PROPOSITION 1.2. *Assume n vertices are drawn independently, on a compact space $\mathcal{K}$, according to a density f such that $\tilde{f}$ is bounded on $\mathcal{K}$. Denote by $l_{TSP}(k_n, n)$ the length of the k_n-TSP on these n vertices, where $2 \le k_n = o(n)$. There exists a universal constant $C > 0$ such that*

$$\limsup_{n\to\infty} \mathbb{E}[l_{TSP}(k_n, n)] \frac{(\|\tilde{f}\|_\infty n)^{\frac{1}{2}\left(1+\frac{1}{k_n-1}\right)}}{k_n - 1} \psi_n \le C.$$

where $\psi_n = 1$ if $k_n \to \infty$ and for any sequence $\psi_n \to 0$ otherwise.

Proof. We first recall that in the Lebesgue differentiation theorem, we can extend the family of balls centered at each point by families of sets with bounded eccentricity $\mathcal{V}$ in other words, there exists $c > 0$ such that every set $U \in \mathcal{V}$ is contained in a ball B with $|U| \ge c|B|$, and such that every point x is contained in arbitrarily small sets of the family $\mathcal{V}$. For instance, in this proof, we can define $\mathcal{V}$ as the family of cubes. The Lebesgue differential theorem gives

$$\tilde{f}(x) = \lim_{U\to x, U\in\mathcal{V}} \frac{1}{|U|}\int_U f,$$

where $U \to x$ means that the sets shrink to x i.e. $x \in U$ and their diameters tend to 0.

Now let $\varepsilon > 0$ be an error tolerance. Consider a cube U^ε such that

$$\left|\frac{1}{|U^\varepsilon|}\int_{U^\varepsilon} f - \|\tilde{f}\|_\infty\right| \le \varepsilon\|\tilde{f}\|_\infty.$$

For convenience, let us write $f(U^\varepsilon) := \frac{1}{|U^\varepsilon|}\int_{U^\varepsilon} f$, and let $N(U^\varepsilon)$ denote the number of vertices contained in U^ε. According to the Hoeffding inequality, with probability $1 - e^{-2\varepsilon^2 f(U^\varepsilon)^2 n}$, U^ε contains at least $n_{U^\varepsilon} := |U^\varepsilon| f(U^\varepsilon)(1-\varepsilon)n \ge |U^\varepsilon|\|\tilde{f}\|_\infty(1-\varepsilon)^2 n$ vertices. We call E_0 this event. Note that $k = o(n_{U^\varepsilon})$. First suppose $k \le n^{1/3}$. Conditionally on E_0, these n_{U^ε} vertices are drawn independently according to a density $\frac{f}{|U^\varepsilon| f(U^\varepsilon)}$ on U^ε. We will now focus on the $k-$TSP in U^ε, which will serve as upper bound for the $k-$TSP on $\mathcal{K}$. From here, the proof is very similar to that of the continuous density case. Let us fix $\alpha > 0$. We start by partitioning U^ε into $P_\alpha := m_\alpha^2$ sub-squares of equal size $\frac{\sqrt{|U^\varepsilon|}}{m_\alpha} \times \frac{\sqrt{|U^\varepsilon|}}{m_\alpha}$ where $m_\alpha := \left\lfloor \frac{1}{\alpha}\sqrt{\frac{{n_{U^\varepsilon}}^{1+\frac{1}{k-1}}}{k-1}} \right\rfloor$. We will show that with high probability, there exists at least one of these sub-squares that contains at least k vertices. Define X_i^α as the number of vertices in sub-square i. Conditionally on E_0, $(X_1^\alpha, \cdots, X_{P_\alpha}^\alpha)$ follows a multinomial where the probability corresponding to sub-square Q_i^α is $p_i = \frac{1}{|U^\varepsilon| f(U^\varepsilon)}\int_{Q_i^\alpha} f \le \frac{|Q_i^\alpha|\|\tilde{f}\|_\infty}{|U^\varepsilon| f(U^\varepsilon)} \le \frac{1}{P_\alpha(1-\varepsilon)}$. Now denote by $A_i^\alpha := \{X_i^\alpha \ge k\}$ the event that sub-square i contains at least k vertices. We first give a lower bound on $\mathbb{P}(A_i^\alpha)$:

$$\mathbb{P}(A_i^\alpha) \ge \binom{n_{U^\varepsilon}}{k} p_i^k \left(1 - \frac{1}{P_\alpha(1-\varepsilon)}\right)^{n_{U^\varepsilon}-k} \ge \frac{n_{U^\varepsilon}^k}{k!}\cdot(1+o(1))\cdot p_i^k.$$

By Jensen's inequality, $\frac{1}{P^\alpha}\sum_{i=1}^{P_\alpha} p_i^k \le \frac{1}{P_\alpha^k}$. Therefore,

$$\sum_{i=1}^{P_\alpha} \mathbb{P}(A_i^\alpha) \ge c\cdot\alpha^{2k-2}\cdot(1+o(1)) \ge \tilde{c}\cdot\alpha^{2k-2}$$

for some constant $\tilde{c} > 0$, so we can use the same proof as in the case of uniform probabilities in the original paper. Then, if $l_{U^\varepsilon}(k, n_{U^\varepsilon})$ denotes the length of the $k-$TSP on the n_{U^ε} vertices in U^ε, we obtain,

$$\mathbb{E}[l_{U^\varepsilon}(k, n_{U^\varepsilon})|E_0] \leq \hat{C}\frac{k-1}{n_{U^\varepsilon}^{\frac{1}{2}\left(1+\frac{1}{k-1}\right)}}\sqrt{|U^\varepsilon|} \leq |U^\varepsilon|^{\frac{-1}{2(k_n-1)}}\frac{C_1(k_n-1)}{(\|\tilde{f}\|_\infty n)^{\frac{1}{2}\left(1+\frac{1}{k_n-1}\right)}},$$

for some constant C_1. If E_0 is not realized, we can use the naive bound $l_{TSP}(k_n, n) \leq l_{TSP}(n, n) \leq 2\sqrt{n} + C$. Therefore,

$$\begin{aligned}\mathbb{E}[l_{TSP}(k_n, n)] &\leq \mathbb{E}[l_{U^\varepsilon}(k, n_{U^\varepsilon})|E_0] + (2\sqrt{n} + C)(1 - \mathbb{P}(E_0)) \\ &\leq |U^\varepsilon|^{\frac{-1}{2(k_n-1)}}\frac{C_1(k_n-1)}{(\|\tilde{f}\|_\infty n)^{\frac{1}{2}\left(1+\frac{1}{k_n-1}\right)}} \cdot (1 + o(1)).\end{aligned}$$

Note that $|U^\varepsilon|^{1/(2(k_n-1))} \to 1$ if $k_n \to \infty$. Otherwise, we can use $|U^\varepsilon|^{1/(2(k_n-1))} \leq |U^\varepsilon|^{-1/2} = o(\psi_n)$ for any sequence $\psi_n \to \infty$ which ends the proof for $k_n \leq n^{1/3}$. In the case where $k \geq n^{1/3}$, the same proof as in the uniform density case shows that

$$\mathbb{E}[l_{U^\varepsilon}(k, n_{U^\varepsilon})|E_0] \leq \sqrt{|U^\varepsilon|}\frac{k-1}{n_{U^\varepsilon}}(2\sqrt{n_{U^\varepsilon}} + C) \leq 2\frac{k-1}{(\|\tilde{f}\|_\infty n)^{\frac{1}{2}\left(1+\frac{1}{k_n-1}\right)}}(1 + o(1)).$$

The proof follows from the same arguments as in the case $k_n \leq n^{1/3}$. □

For the case $k = \Theta(n)$, we expect a constant-factor approximation for the $k-$TSP to perform the TSP on a set with maximal average density and area $\Theta(k/n)$. In the following, we state this generalization as a claim without proof. A possible proof sketch would use similar techniques to the analysis developed for the TRP in the original paper.

Claim 1.3. *Assume n vertices are drawn independently, on a compact space $\mathcal{K}$, according to a density f. Let $\varepsilon > 0$. Denote by $l_{TSP}(k_n, n)$ the length of the k_n-TSP on these n vertices, where $\varepsilon n \leq k_n \leq n$. There exists constants $0 < c_\varepsilon < C$ such that*

$$c_\varepsilon \leq \liminf_{n\to\infty}\frac{\mathbb{E}[l_{TSP}(k_n, n)]}{\sqrt{n}g_f(k_n/n)} \leq \limsup_{n\to\infty}\frac{\mathbb{E}[l_{TSP}(k_n, n)]}{\sqrt{n}g_f(k_n/n)} \leq C,$$

where if we denote by F the cumulative distribution of f and $y_0 = \inf\{y : 1 - F(y) \leq k_n/n\}$,

$$g_f(k_n/n) = \int \sqrt{f}\mathbf{1}_{f>y_0} + \frac{k_n/n - (1 - F(y_0))}{\sqrt{y_0}}.$$

2 The Ψ-TRP

In the main paper, we analyzed the TRP under fairness considerations. In particular, we showed that achieving efficency while ensuring max-min fairness asymptotically is possible. Here, we propose another notion of fairness and give similar positive results. Recall that the TRP objective of a given tour is

$$\sum_{i=1}^{n} l_i,$$

where l_i is the latency at vertex i. In resource allocation problems, this objective corresponds to the utilitarian principle i.e. maximizing the total utility. A common approach to fairness consists of maximizing $\sum_j f(u_j)$ where f is a concave function and u_j denotes the utility of player j. In particular, the log function yields the proportional fairness solution under mild convexity assumptions [1]. We adapt this idea to our setting by changing the latency objective. Specifically, for any increasing function Ψ, we can define the $\Psi-$TRP, which seeks a tour that minimizes the objective:

$$\sum_{i=1}^{n} \Psi(l_i).$$

To capture fairness considerations, we assume that Ψ is convex. We show that, for a large class of functions Ψ, our approximation algorithm for the TRP is also constant-factor optimal for the Ψ-TRP, hence encapsulating this notion of fairness. Indeed, our analysis for the TRP generalizes to the $\Psi-$TRP when Ψ is a convex monomial. This is formalized in the following proposition, which we prove in the next sections.

Proposition 2.1. *Assume all n vertices are drawn according to a distribution with density f on a compact space $\mathcal{K} \subset \mathbb{R}^2$. Let $\alpha \geq 1$ and $\Psi_\alpha : x \mapsto x^\alpha$ the power function. Denote by $l_{\alpha-TRP}$ the optimal $\Psi_\alpha-TRP$ objective of a tour for the $\Psi_\alpha-TRP$. Then,*

$$c_\alpha \int_{\mathcal{K}} g_\alpha(f,x)dx \leq \liminf_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \leq \limsup_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \leq C_\alpha \int_{\mathcal{K}} g_\alpha(f,x)dx$$

where $0 < c_\alpha < C_\alpha$ are two constants depending only in α and

$$g_\alpha(f,x) = \begin{cases} f(x)\left(\int_{\mathcal{K}} \sqrt{f}\cdot \mathbf{1}_{f<f(x)}\right)^\alpha, & \text{if } \int_{\mathcal{K}} \mathbf{1}_{f=f(x)} = 0 \\ \sqrt{f(x)} \frac{\left(\int_{\mathcal{K}} \sqrt{f}\cdot \mathbf{1}_{f\leq f(x)}\right)^{\alpha+1} - \left(\int_{\mathcal{K}} \sqrt{f}\cdot \mathbf{1}_{f<f(x)}\right)^{\alpha+1}}{(\alpha+1)\int_{\mathcal{K}} \mathbf{1}_{f=f(x)}}, & \text{otherwise.} \end{cases}$$

We use similar proof ideas as for the probabilistic bounds of the classical TRP. However, because Ψ_α is non-linear for $\alpha > 1$, the arguments are more technical. In particular for the lower bound, we divide the tour into sub-paths in each sub-square of the partition but with the additional constraint that all sub-paths should visit the same number of vertices. The non-linearity of Ψ_α also affects the form of the integrand $g_\alpha(f,\cdot)$ for degenerate levels of the density function when $\int_{\mathcal{K}} \mathbf{1}_{f=f(x)} > 0$. As a result, the proof of convergence of the integral of $g_\alpha(\phi,\cdot)$ to the integral of $g_\alpha(f,\cdot)$, for fine piece-wise constant approximations ϕ of f, is more technical than the equivalent result for the TRP.

Furthermore, the upper bound is reached by the same approximating scheme as for the TRP in which we partition the space in sub-squares and visit sub-squares by decreasing order of density. In particular, this scheme is also constant-factor optimal for the $\Psi-$TRP. Using the same arguments, we can generalize Proposition 2.1 to any linear combination of monomials where the leading term is a of the form $x \mapsto c\cdot x^\alpha$ where $c > 0$ and $\alpha \geq 1$. In other words, the competitive ratio between the fairness-maximizing $\Psi-$TRP and the efficiency-maximizing TRP is asymptotically 1.

2.1 A lower bound

Proposition 2.2. *Assume all n vertices are drawn according to a distribution with density f on a compact space $\mathcal{K} \subset \mathbb{R}^2$. Let $\alpha \geq 1$ and $\Psi_\alpha : x \mapsto x^\alpha$ the power function. Denote by*

$l_{\alpha-TRP}$ *the optimal TRP objective of a tour for the* $\Psi_\alpha-TRP$*. Then,*

$$\liminf_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \geq c_\alpha \int_{\mathcal{K}} g_\alpha(f,x)dx,$$

where $c_\alpha := \frac{1}{(\pi e)^{\alpha/2}}$ *is a constant and*

$$g_\alpha(f,x) = \begin{cases} f(x)\left(\int_{\mathcal{K}} \sqrt{f}\cdot \mathbf{1}_{f<f(x)}\right)^{\alpha}, & \text{if } \int_{\mathcal{K}} \mathbf{1}_{f=f(x)} = 0 \\ \sqrt{f(x)}\frac{\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f\leq f(x)}\right)^{\alpha+1} - \left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha+1}}{(\alpha+1)\int_{\mathcal{K}}\mathbf{1}_{f=f(x)}}, & \text{otherwise,} \end{cases}$$

is a function that depends only on α *and* f*.*

Proof. We take the same notations as in the proof of the lower bound of Theorem 3 from [2]. Again, we first start by the case where f has support in the unit square $[0,1]^2$ and has the form

$$f = \sum_{k=1}^{m^2} f_k \mathbf{1}_{Q_k},$$

where $\{Q_i\}$ is the regular partition of the unit square into m^2 sub-squares. We define the margin

$$\mathcal{M} = \bigcup_{1\leq k\leq m^2} Q_k \cap \left(\partial Q_k + \varepsilon_m^{(k)} B(0,1)\right),$$

for $\varepsilon_m^{(k)} := \frac{\varepsilon}{m}\sqrt{\frac{f_*}{f_k}}$. Note that this is a smaller margin than what was considered in the proof of the lower bound of Theorem 3 from [2]. We can have estimates for the number of vertices in the margin similar to Lemma 3 of [2]. Finally, we define the event E_0 in which for all $1 \leq k \leq m^2$ such that $f_k > 0$,

$$\frac{f_k}{2m^2}n \leq N_k \leq \frac{3f_k}{2m^2}n, \quad l_{TSP(Q_k)}\left(\left\lceil \varepsilon\cdot e\sqrt{\pi\frac{3f_*}{2m^2}n}\right\rceil, N_k\right) > \varepsilon_n^{(k)}.$$

Let us estimate the probability of the event E_0. By the proof of Lemma 4 of [2],

$$\mathbb{P}\left[\left|N_k - \frac{f_k}{m^2}n\right| \geq \frac{f_k}{2m^2}n\right] \leq 2e^{-\frac{1}{12}\cdot\frac{f_k}{m^2}n}.$$

We now use Corollary 1 of [2] to each of the sub-squares. For $1 \leq k \leq m^2$, such that $f_k > 0$,

$$\begin{aligned} \mathbb{P}\left[l_{TSP(Q_k)}\left(\left\lceil \varepsilon\cdot e\sqrt{\pi\frac{3f_*}{2m^2}n}\right\rceil, N_k\right) \leq \varepsilon_m^{(k)} \,\middle|\, \frac{f_k}{2m^2}n < N_k < \frac{3f_k}{2m^2}n\right] \\ \leq \mathbb{P}\left[l_{TSP(Q_k)}\left(\left\lceil \varepsilon\cdot e\sqrt{\pi\frac{3f_*}{2m^2}n}\right\rceil, \frac{3f_k}{2m^2}n\right) \leq \varepsilon_m^{(k)}\right] \\ = o\left(e^{-\varepsilon\cdot e\sqrt{\pi\frac{3f_* n}{2m^2}}}\right). \end{aligned}$$

Finally, the probability of E_0 is $1 - o\left(e^{-c\varepsilon\sqrt{\frac{f_* n}{m^2}}}\right)$ for some constant $c > 0$.

In the next steps we will assume that this event is met. We are now ready to use an equivalent of Lemma 5 from [2] to each sub-path in Q_k which is not completely included in the margin. However, we will need all sub-paths to visit same number of vertices. Denote by $l_{\Psi-TRP}$ the optimal objective and consider an optimal tour. We order the sub-paths $\mathcal{P}_1, \cdots \mathcal{P}_P$ which are not completely included in the margin. Also, we denote by $k(i)$ the index of the sub-square containing $\mathcal{P}_i$. We divide $\mathcal{P}_i$ into smaller sub-paths of length exactly $n_* = \left\lceil \varepsilon \cdot e\sqrt{\pi \frac{3f_*}{2m^2} n} \right\rceil$. Since the number of vertices visited by $\mathcal{P}_i$ might not be a multiple of n_*, some vertices will be left out. For any path $\mathcal{P}_i$, if $n(\mathcal{P}_i) \geq n_*$, then at most $n(\mathcal{P}_i)/2$ vertices will be left out. We will denote $\mathcal{P}_i^1, \cdots \mathcal{P}_i^{t_i}$ the corresponding created sub-paths containing exactly n_* edges. Note that $t_i = \left\lfloor \frac{n(\mathcal{P}_i)}{n_*} \right\rfloor$. We now treat paths with $n(\mathcal{P}_i) < n_*$ separately, which we will call low-density paths. Let L be the set of indices of low density paths. For a given low-density path $\mathcal{P}_i$, we artificially add $n_* - n(\mathcal{P}_i)$ vertices from later low-density paths $\mathcal{P}_j$ in the same sub-square as $\mathcal{P}_i$, where $j > i$. At the end of this process, at most one low-density path remains, which we will leave out. Let us denote by $\hat{\mathcal{P}}_i$ for $i \in \tilde{L}$, the corresponding constructed paths from low-density paths. Note that we have $\tilde{L} \subset L$, but not necessarily an equality because the process can potentially remove all vertices of some low-density paths. In the following, if $\mathcal{P}$ is a sub-path, we will denote by $l(\mathcal{P})$ its length. Let us summarize the obtained lower bound.

$$\begin{aligned}
l_{\Psi-TRP} &= \sum_{v \in V} \Psi_\alpha(\text{completion time of } v) \\
&= \sum_{1 \leq i \leq P} \sum_{v \in \mathcal{P}_i} \Psi_\alpha(\text{completion time of } v) \\
&\geq \sum_{i \notin L} \sum_{t=1}^{t_i} \sum_{v \in \mathcal{P}_i^t} \Psi_\alpha(\text{completion time of } v) + \sum_{i \in \tilde{L}} \sum_{v \in \hat{\mathcal{P}}_i} \Psi_\alpha(\text{completion time of } v) \\
&\geq \sum_{i \notin L} \sum_{t=1}^{t_i} \sum_{v \in \mathcal{P}_i^t} \Psi_\alpha\left(\sum_{j=1}^{i-1} l(\mathcal{P}_j) + \sum_{u=1}^{t-1} l(\mathcal{P}_i^u)\right) + \sum_{i \in \tilde{L}} \sum_{v \in \hat{\mathcal{P}}_i} \Psi_\alpha\left(\sum_{j=1}^{i-1} l(\mathcal{P}_j)\right) \\
&= n_* \left[\sum_{i \notin L} \sum_{t=1}^{t_i} \Psi_\alpha\left(\sum_{j=1}^{i-1} l(\mathcal{P}_j) + \sum_{u=1}^{t-1} l(\mathcal{P}_i^u)\right) + \sum_{i \in \tilde{L}} \Psi_\alpha\left(\sum_{j=1}^{i-1} l(\mathcal{P}_j)\right)\right] \\
&\geq n_* \sum_{1 \leq i \leq \tilde{P}} \Psi_\alpha\left(\sum_{j=1}^{i-1} l(\tilde{\mathcal{P}}_j)\right),
\end{aligned}$$

where we have listed the new sub-paths containing n_* vertices: $\tilde{\mathcal{P}}_1, \cdots \tilde{\mathcal{P}}_{\tilde{P}}$ with the order given by the original tour — the ordering where we omit added vertices to low-density sub-paths. The length of the subpath $\tilde{\mathcal{P}}_i$ is the length of the corresponding subpath $\tilde{\mathcal{P}}_j^t$ if it came from a non low-density path. Otherwise, we define it as $l(\tilde{\mathcal{P}}_i) := l(\mathcal{P}_j)$ where $\tilde{\mathcal{P}}_i = \hat{\mathcal{P}}_j$. This corresponds to lower bounding the contribution of added vertices in low-density sub-paths, to the $\Psi - TRP$ objective. A key observation is that we can have a similar result to that of Lemma 5 from [2]. Again, we will denote by $k(i)$ the index of the sub-square containing sub-path $\tilde{\mathcal{P}}_i$, i.e. $\tilde{\mathcal{P}}_i \subset Q_{k(i)}$.

Lemma 2.3. *Let $1 \le i \le \tilde{P}$. Under the event E_0, for n sufficiently large, we can give the lower bound*

$$l(\tilde{\mathcal{P}}_i) \ge \frac{n_*}{2\sqrt{\pi e \cdot f_{k(i)} n}}.$$

Proof. Under E_0, no path containing at n_* vertices has lower length than $\varepsilon_m^{(k)}$. Let us first consider the case of a sub-path $\tilde{p}_i$ corresponding to a sub-path of a non low-density sub-path p_j. Then, $\tilde{p}_i$ is a "true" sub-path of the original tour and contains n_* vertices. Therefore, $l_{\tilde{p}_i} \ge \varepsilon_m^{(k(i))}$. Let us now consider a sub-path $\tilde{p}_i$ corresponding to a low-density sub-path p_j for $j \in L$. Recall that p_j is a sub-path of $Q_{k(i)}$ which is not entirely contained in the margin. Therefore, $l_{p_j} \ge \varepsilon_m^{(k(i))}$. In summary, for all $1 \le i \le \tilde{P}$,

$$l_{\tilde{p}_i} \ge \varepsilon_m^{(k(i))} \ge \frac{n_*}{2\sqrt{\pi e \cdot f_{k(i)} n}},$$

where the second inequality is true for n sufficiently large. □

Therefore, under E_0 we have the following lower bound,

$$\begin{aligned}
l_{\Psi-TRP} &\ge n_* \sum_{1\le i\le \tilde{P}} \Psi_\alpha\left(\sum_{j=1}^{i-1} l_{\tilde{p}_j}\right) \\
&\ge n_* \sum_{1\le i\le \tilde{P}} \Psi_\alpha\left(\frac{1}{2\sqrt{\pi e n}} \sum_{j=1}^{i-1} \frac{n_*}{\sqrt{f_{k(j)}}}\right) \\
&= \frac{n_*^{\alpha+1}}{2^\alpha (\pi e n)^{\alpha/2}} \sum_{1\le i\le \tilde{P}} \Psi_\alpha\left(\sum_{j=1}^{i-1} \frac{1}{\sqrt{f_{k(j)}}}\right) \\
&\ge \frac{n_*^{\alpha+1}}{2^\alpha (\pi e n)^{\alpha/2}} \cdot \min_{\sigma\in\mathcal{S}_{\tilde{P}}} \sum_i \Psi_\alpha\left(\sum_{j=1}^{i-1} \frac{1}{\sqrt{f_{k(\sigma(j))}}}\right).
\end{aligned}$$

Let us now give an equivalent of Lemma 6 from [2].

Lemma 2.4. *The minimum objective of the optimization problem*

$$\min_{\sigma\in\mathcal{S}_{\tilde{P}}} \sum_i \Psi_\alpha\left(\sum_{j=1}^{i-1} \frac{1}{\sqrt{f_{k(\sigma(j))}}}\right).$$

is given by ordering sub-paths by increasing order of $\frac{1}{\sqrt{f_{k(i)}}}$, i.e. decreasing order of $f_{k(i)}$.

Proof. In this proof, we will denote by C_σ the objective of the minimization problem for $\sigma \in \mathcal{S}_P$, i.e.

$$C_\sigma := \min_{\sigma\in\mathcal{S}_{\tilde{P}}} \sum_i \Psi_\alpha \left(\sum_{j=1}^{i-1} \frac{1}{\sqrt{f_{k(\sigma(j))}}} \right).$$

Let $1 \leq i < \tilde{P}$. We will compare C_σ and $C_{\tilde{\sigma}}$ where $\tilde{\sigma}$ is the permutation σ but the $i-$th and $(i+1)-$th index are interchanged:

$$\tilde{\sigma}(r) = \begin{cases} \sigma(r), & r \notin \{i, i+1\} \\ \sigma(i+1), & r = i \\ \sigma(i), & r = i+1. \end{cases}$$

Then,

$$C_{\tilde{\sigma}} - C_\sigma = \Psi\left(\frac{1}{\sqrt{f_{k(\sigma(i+1))}}} + \eta \right) - \Psi\left(\frac{1}{\sqrt{f_{k(\sigma(i))}}} + \eta \right),$$

where $\eta = \sum_{t<i} \frac{n_{p_{\sigma(t)}}}{\sqrt{f_{k(\sigma(t))}}}$. Assume that we have $\frac{1}{\sqrt{f_{k(\sigma(i+1))}}} \leq \frac{1}{\sqrt{f_{k(\sigma(i))}}}$. Then, the objective is decreases when we place $\sigma(i+1)$ in $i-$th position: $C_{\tilde{\sigma}} \leq C_\sigma$. We then use this argument to order sequentially the permutation σ by decreasing order of $f_{k(i)}$. This ends the proof of the lemma. □

Let us now give estimates on the right hand of the inequality. Denote by σ^* the ordering on the sub-squares Q_k such that $\frac{1}{\sqrt{f_{\sigma^*(k)}}}$ is increasing in k. Then under E_0,

$$\begin{aligned}
&\min_{\sigma\in\mathcal{S}_{\tilde{P}}} \sum_i n_* \Psi_\alpha \left(\sum_{j=1}^{i-1} \frac{n_*}{\sqrt{f_{k(\sigma(j))}}} \right) \\
&\quad \geq \sum_{1\leq k\leq m^2} \left(\sum_{i,\, \tilde{p}_i \in Q_{\sigma^*(k)}} n_* \right) \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1} \frac{1}{\sqrt{f_{\sigma^*(t)}}} \left(\sum_{i,\, \tilde{p}_i\in Q_{\sigma^*(t)}} n_* \right) \right] \\
&\quad \geq \sum_{1\leq k\leq m^2} \frac{N_{\sigma^*(k)} - |V\cap Q_{\sigma^*(k)} \cap \mathcal{M}| - n_*}{2} \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1} \frac{N_{\sigma^*(t)} - |V \cap Q_{\sigma^*(t)} \cap \mathcal{M}| - n_*}{2\sqrt{f_{\sigma^*(t)}}} \right] \\
&\quad \geq \frac{1}{2^{\alpha+1}} \sum_{1\leq k\leq m^2} N_{\sigma^*(k)} \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1} \frac{N_{\sigma^*(t)}}{\sqrt{f_{\sigma^*(t)}}} \right] - \frac{1}{2^{\alpha+1}} \sum_{1\leq k\leq m^2} (|V \cap Q_{\sigma^*(k)} \cap \mathcal{M}| + n_*) \cdot \Psi_\alpha \left[\frac{n}{\sqrt{f_*}} \right] \\
&\qquad - \frac{\alpha}{2^{\alpha+1}} \sum_{1\leq k\leq m^2} N_{\sigma^*(k)} \sum_{t=1}^{k-1} \frac{|V \cap Q_{\sigma^*(t)} \cap \mathcal{M}| + n_*}{\sqrt{f_*}} \cdot \Psi_{\alpha-1} \left[\frac{n}{\sqrt{f_*}} \right] \\
&\quad \geq \frac{n^{\alpha+1}}{4^{\alpha+1} m^{2\alpha+2}} \sum_{1\leq k\leq m^2} f_{\sigma^*(k)} \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1} \sqrt{f_{\sigma^*(t)}} \right] - \frac{|V\cap \mathcal{M}| + m^2 n_*}{2^{\alpha+1}} \cdot \Psi_\alpha \left[\frac{n}{\sqrt{f_*}} \right] \\
&\qquad - \frac{\alpha \cdot n}{2^{\alpha+1}} \frac{|V \cap \mathcal{M}| + m^2 n_*}{\sqrt{f_*}} \cdot \Psi_{\alpha-1} \left[\frac{n}{\sqrt{f_*}} \right]
\end{aligned}$$

$$= \frac{n^{\alpha+1}}{4^{\alpha+1}m^{2\alpha+2}} \sum_{1\le k\le m^2} f_{\sigma^*(k)} \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1} \sqrt{f_{\sigma^*(t)}}\right] - \frac{n^\alpha(1+\alpha)}{2^{\alpha+1} f_*^{\alpha/2}}(|V\cap\mathcal{M}| + m^2 n_*).$$

Using Lemma 3 of [2], we obtain that with probability $1-o(e^{-c\varepsilon\sqrt{\frac{f_* n}{m^2}}})$, the event E_0 is met and $|V\cap\mathcal{M}|\le 8\varepsilon n$. Therefore, we can take $\varepsilon>0$ sufficiently small so that

$$l_{TRP} \ge \frac{1+o(1)}{8^{1+\alpha}(\pi e)^{\alpha/2} m^{2\alpha+2}} n^{1+\alpha/2} \sum_{1\le k\le m^2} f_{\sigma^*(k)} \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1}\sqrt{f_{\sigma^*(t)}}\right].$$

Define a new constant $c_\alpha := \frac{1}{8^{1+\alpha}(\pi e)^{\alpha/2}}$, we now obtain the desired result.

$$\begin{aligned}
\liminf_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} &\ge \liminf_{n\to\infty} \mathbb{P}(E_0)\cdot \frac{\mathbb{E}[l_{\Psi-TRP}|E_0]}{n^{1+\alpha/2}} \\
&\ge \liminf_{n\to\infty}\left(1-o\left(e^{-c\varepsilon\sqrt{\frac{f_* n}{m^2}}}\right)\right)\frac{c_\alpha(1+o(1))}{m^{2\alpha+2}} \sum_{1\le k\le m^2} f_{\sigma^*(k)}\Psi_\alpha\left[\sum_{t=1}^{k-1}\sqrt{f_{\sigma^*(t)}}\right] \\
&\ge \frac{c_\alpha}{m^{2\alpha+2}} \sum_{1\le k\le m^2} f_{\sigma^*(k)}\cdot\Psi_\alpha\left[\sum_{t=1}^{k-1}\sqrt{f_{\sigma^*(t)}}\right].
\end{aligned}$$

We will now make the link between the discrete sum and the integral formula. To do so, we aggregate sub-squares who have same density f_k. If $f^1>\cdots>f^S$ are values taken by the density function, We obtain a partition $\{1,\cdots,m^2\}=\bigcup_{1\le s\le S} A_s$, where $A_s=\{k:\ f_k=f^s\}$ contains the indices of sub-squares having density f^s. Note that because the values $f^1,\cdots,f^S$ are ordered, so are the sets A_s i.e. all elements of A_2 are larger than elements of A_1, etc. Then,

$$\begin{aligned}
&\frac{1}{m^{2\alpha+2}} \sum_{1\le k\le m^2} f_{\sigma^*(k)}\cdot\Psi_\alpha\left[\sum_{t=1}^{k-1}\sqrt{f_{\sigma^*(t)}}\right] \\
&= \sum_{1\le k\le m^2} \int_{Q_{\sigma^*(k)}} f(x)\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{Q_{\sigma^*(1)}\cup\cdots\cup Q_{\sigma^*(k-1)}}\right)^\alpha dx \\
&= \sum_{s=1}^{S}\sum_{k\in A_s}\int_{Q_{\sigma^*(k)}} f(x)\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)} + \sqrt{f(x)}\cdot\mathbf{1}_{\bigcup_{l\in A_s, l<k} Q_{\sigma^*(k)}}\right)^\alpha dx.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sum_{s=1}^{S}\int_0^{\mathcal{A}(\bigcup_{l\in A_s} Q_{\sigma^*(k)})} f(x)\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)} + t\sqrt{f(x)}\right)^\alpha dt - \frac{1}{m^{2\alpha+2}}\sum_{1\le k\le m^2} f_{\sigma^*(k)}\cdot\Psi_\alpha\left[\sum_{t=1}^{k-1}\sqrt{f_{\sigma^*(t)}}\right] \\
&= \sum_{s=1}^{S}\sum_{k\in A_s} f(x)\int_0^{\mathcal{A}(Q_{\sigma^*(k)})}\left[\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)} + \sqrt{f(x)}\cdot\mathcal{A}(\cup_{l\in A_s,l<k}Q_{\sigma^*(k)}) + t\sqrt{f(x)}\right)^\alpha\right. \\
&\qquad \left. - \left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)} + \sqrt{f(x)}\cdot\mathcal{A}(\cup_{l\in A_s,l<k}Q_{\sigma^*(k)})\right)^\alpha\right] dt
\end{aligned}$$

$$
\begin{aligned}
&\le \sum_{s=1}^{S} \sum_{k \in A_s} f(x) \int_0^{\mathcal{A}(Q_{\sigma^*(k)})} \alpha \left(\int_{\mathcal{K}} \sqrt{f} \right)^{\alpha-1} t\sqrt{f(x)} dt \\
&= \frac{\alpha}{2m^2} \left(\int_{\mathcal{K}} \sqrt{f} \right)^{\alpha-1} \int_{\mathcal{K}} f^{3/2}.
\end{aligned}
$$

Also note that

$$
\sum_{s=1}^{S} \int_0^{\mathcal{A}\left(\bigcup_{l \in A_s} Q_{\sigma^*(k)}\right)} f(x) \left(\int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{f<f(x)} + t\sqrt{f(x)} \right)^{\alpha} dt = \int_{\mathcal{K}} g_\alpha(f,x) dx.
$$

Finally, we have

$$
\liminf_{n \to \infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \ge c_\alpha \int_{\mathcal{K}} g_\alpha(f,x)dx - \frac{c_\alpha \alpha}{2m^2} \left(\int_{\mathcal{K}} \sqrt{f} \right)^{\alpha-1} \int_{\mathcal{K}} f^{3/2}.
$$

We can repeat the same procedure with a finest partition of the unit square $[0,1]^2$ into $(\beta m)^2$ sub-squares where $\beta \in \mathbb{N}^*$. For β sufficiently large, we obtain

$$
\frac{\alpha}{2m^2} \left(\int_{\mathcal{K}} \sqrt{f} \right)^{\alpha-1} \int_{\mathcal{K}} f^{3/2} \le \frac{1}{2} \int_{\mathcal{K}} g_\alpha(f,x)dx.
$$

Then, with this partition we obtain the desired result

$$
\liminf_{n \to \infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \ge \tilde{c}_\alpha \int_{\mathcal{K}} g_\alpha(f,x)dx,
$$

where $\tilde{c}_\alpha = \frac{c_\alpha}{2(\alpha+1)} = \frac{1}{2 \cdot 8^{\alpha+1} (\pi e)^{\alpha/2} (\alpha+1)}$. This ends the proof for the densities of the form

$$
f(x) = \sum_{k=1}^{m^2} f_k \mathbf{1}_{Q_k}(x).
$$

Note that with the same proof, we can tighten the constant $\tilde{c}_\alpha$ to be $\frac{1}{(\pi e)^{\alpha/2}(\alpha+1)}$.

We now turn to general distributions with continuous densities. To do so, we need an equivalent of Lemma 7 from [2], which is given by Lemma 2.6. Similarly to the proof of the lower bound of Theorem 3 of [2], let us now consider the general case of an absolutely continuous density f on a compact space $\mathcal{K}$. By a scaling argument, we can suppose without loss of generality that $\mathcal{K} \subset [0,1]^2$. For any $\varepsilon > 0$, we use Lemma 2.6 to take a density ϕ of the same form as above

$$
\phi(x) = \sum_{k=1}^{m^2} \phi_k \mathbf{1}_{Q_k}(x),
$$

such that $\|\phi - f\|_\infty \le \varepsilon$ and $\left| \int_{\mathcal{K}} g_\alpha(\phi) - g_\alpha(f) \right| \le \varepsilon$. By a coupling argument, we can construct a joint distribution (X,Y) such that X (resp. Y) has density f (resp. ϕ), and $\mathbb{P}(X \neq Y) \le 2\int_{\mathcal{K}} |\phi(x) - f(x)|dx \le 2\varepsilon$. On the event $\{X_i = Y_i, 1 \le i \le n\}$, the $\Psi-$TRP lengths coincide. Therefore, we can use the estimates on ϕ to show that

$$
\liminf_{n \to \infty} \frac{\mathbb{E}[l_{TRP}(f)]}{n^{\alpha/2}} \ge (1-2\varepsilon)^n c_\alpha \int_{\mathcal{K}} g_\alpha(\phi)
$$

$$\geq (1-2\varepsilon)^n c_\alpha \left(\int_{\mathcal{K}} g_\alpha(f) - \varepsilon \right).$$

Since this is valid for any $\varepsilon > 0$, the desired result follows.

$$\liminf_{n\to\infty} \frac{\mathbb{E}[l_{TRP}(f)]}{n^{\alpha/2}} \geq c_\alpha \int_{\mathcal{K}} g_\alpha(f)$$

This ends the proof of the Proposition. ☐

2.2 An upper bound

We now give a constructive proof of an upper bound. The resulting constructed tour is constant-factor from the optimal $\Psi-$TRP tour.

PROPOSITION 2.5. *Assume all n vertices are drawn according to a distribution with density f on a compact space $\mathcal{K} \subset \mathbb{R}^2$. Let $\alpha \geq 1$ and $\Psi_\alpha : x \mapsto x^\alpha$ the power function. Denote by $l_{\alpha-TRP}$ the optimal TRP objective of a tour for the $\Psi_\alpha-$TRP. Then,*

$$\liminf_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \leq C_\alpha \int_{\mathcal{K}} g_\alpha(f,x)dx,$$

where $C_\alpha > 0$ is a constant depending only on α. Furthermore, there exists a simple way to construct a tour that achieves the provided upper bound.

Proof. Let $\varepsilon > 0$. Take $m > 0$ and a piece-wise constant density ϕ approximating f, given by Lemma 2.6. Similarly to the tour constructed in the proof of the upper bound of Theorem 3 of [2], if we order the sub-squares by decreasing value of ϕ and denote σ this ordering, our tour will follow a TSP tour on $Q_{\sigma(1)}$, then on $Q_{\sigma(2)}$, until $Q_{\sigma(m^2)}$. We will now show that this tour is constant-factor optimal on the high-event probability E_0 in which

$$\frac{\phi_k}{2m^2} n \leq N_k \leq \frac{3\phi_k}{2m^2} n,$$

for all $1 \leq k \leq m^2$, where N_k is the count of vertices in sub-square Q_k. By the Chernoff bound, $\mathbb{P}(E_0) = 1 - o\left(e^{-c\frac{\phi_* n}{m^2}}\right)$, where $\phi_* := \min\{\phi_k\}$ and $c > 0$ a constant. Let us now analyze the $\Psi-$TRP objective of this tour on E_0. On each sub-square, by the BHH theorem, the length l_{TSP}^k of the optimal TSP satisfies

$$l_{TSP}^k \leq C\beta_{TSP} \sqrt{\frac{3\phi_k n}{2} \frac{1}{m^2}}.$$

for $C > 0$ a constant and any n sufficiently large. Then, if $\hat{l}_{\Psi-TRP}$ denotes the objective of the constructed tour, for n sufficiently large,

$$\begin{aligned}
\hat{l}_{\Psi-TRP} &\leq \sum_{k=1}^{m^2} N_{\sigma(k)} \Psi_\alpha \left(\sum_{l=1}^{k} l_{TSP}^{\sigma(l)} + (k-1)\sqrt{2} \right) \\
&\leq \frac{3n}{2m^2} \sum_{k=1}^{m^2} \phi_{\sigma(k)} \Psi_\alpha \left(C\beta_{TSP} \sqrt{\frac{3n}{2}} \frac{1}{m^2} \sum_{l=1}^{k} \sqrt{\phi_{\sigma(l)}} + (k-1)\sqrt{2} \right)
\end{aligned}$$

$$\leq \left(\frac{3}{2}\right)^{1+\alpha/2} C^\alpha \beta_{TSP}^\alpha \frac{n^{1+\alpha/2}}{m^2} \sum_{k=1}^{m^2} \phi_{\sigma(k)} \left[\Psi_\alpha \left(\frac{1}{m^2}\sum_{l=1}^{k} \sqrt{\phi_{\sigma(l)}}\right) + \frac{2\alpha k}{C\beta_{TSP}\sqrt{3n}} \cdot 2 \left\|\sqrt{\phi}\right\|_1^{\alpha-1}\right]$$
$$\leq \left(\frac{3}{2}\right)^{1+\alpha/2} C^\alpha \beta_{TSP}^\alpha \frac{n^{1+\alpha/2}}{m^{2+2\alpha}} \sum_{k=1}^{m^2} \phi_{\sigma(k)} \Psi_\alpha \left(\sum_{l=1}^{k} \sqrt{\phi_{\sigma(l)}}\right) + \left(\frac{3}{2}\right)^{1+\alpha/2} (C\beta_{TSP})^{\alpha-1} \frac{4\alpha m^2 n^{(\alpha+1)/2}}{\sqrt{3}}.$$

Therefore, with $C_\alpha := \left(\frac{3}{2}\right)^{1+\alpha/2} C^\alpha \beta_{TSP}^\alpha$, we obtain

$$\begin{aligned}
\liminf_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} &\leq \frac{C_\alpha}{m^{2\alpha+2}} \sum_{1\leq k\leq m^2} \phi_{\sigma(k)} \cdot \Psi_\alpha \left[\sum_{t=1}^{k} \sqrt{\phi_{\sigma(t)}}\right] \\
&\leq \frac{C_\alpha}{m^{2\alpha+2}} \sum_{1\leq k\leq m^2} \phi_{\sigma(k)} \cdot \Psi_\alpha \left[\sum_{t=1}^{k-1} \sqrt{\phi_{\sigma(t)}}\right] + \frac{C_\alpha}{m^4} \sum_{1\leq k\leq m^2} \phi_{\sigma(k)} \cdot \alpha \sqrt{\phi_{\sigma(k)}} \\
&\leq C_\alpha \int_{\mathcal{K}} g_\alpha(\phi) + \frac{C_\alpha}{m^2} \int_{\mathcal{K}} \phi^{3/2}.
\end{aligned}$$

We can take a finer subdivision and take m sufficiently large so that finally,

$$\liminf_{n\to\infty} \frac{\mathbb{E}\left[l_{\Psi-TRP}\right]}{n^{1+\alpha/2}} \leq \tilde{C}_\alpha \int_{\mathcal{K}} g_\alpha(\phi),$$

where $\tilde{C}_\alpha = 2C_\alpha$. Note that with the same proof we can get the same result with $\tilde{C}_\alpha = \beta_{TSP}^\alpha$. This ends the proof. ∎

2.3 Technical lemma

Lemma 2.6. *Let f be a density on $\mathcal{K} \subset [0,1]^2$. For any $\varepsilon > 0$, there exists a density ϕ of the form*

$$\phi(x) = \sum_{k=1}^{m^2} \phi_k \mathbf{1}_{Q_k}(x)$$

such that

$$\|\phi - f\|_1 \leq \varepsilon, \quad \textit{and} \quad \left|\int_{\mathcal{K}} g_\alpha(\phi) - \int_{\mathcal{K}} g_\alpha(f)\right| \leq \varepsilon.$$

Proof. Let $\delta > 0$ an error parameter. Similarly to the proof of Lemma 7 from [2], we first take a density ϕ_ε of the right form such that $\varepsilon \leq \delta$ and

$$\|\phi_\varepsilon - f\|_1 \leq \varepsilon, \quad \text{and} \quad \|\sqrt{\phi_\varepsilon} - \sqrt{f}\|_1 \leq \varepsilon.$$

We choose ϕ_ε such that all ϕ_k are distinct. We will now write ϕ instead of ϕ_ε. Again, $\|\sqrt{f}\|_1, \|\sqrt{\phi}\|_1 \leq 1$. We first introduce a new function $\tilde{g}_\phi$ which we will use as intermediary.

$$\tilde{g}_\alpha(\phi, x) = \begin{cases} \phi(x)\left(\int_{\mathcal{K}} \sqrt{\phi} \cdot \mathbf{1}_{f<f(x)}\right)^\alpha & \text{if } \int_{\mathcal{K}} \mathbf{1}_{f=f(x)} = 0, \\ \sqrt{f(x)} \frac{\left(\int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{f\leq f(x)}\right)^{\alpha+1} - \left(\int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{f<f(x)}\right)^{\alpha+1}}{(\alpha+1)\int_{\mathcal{K}} \mathbf{1}_{f=f(x)}} & \text{otherwise,} \end{cases}$$

Let us start by giving an estimate that will later be useful.

$$\left|\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha}-\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha}\right|\leq\int_{\mathcal{K}}\left|\sqrt{\phi}-\sqrt{f}\right|\cdot\mathbf{1}_{f<f(x)}$$
$$\cdot\,\alpha\max\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{f<f(x)},\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha-1}$$
$$\leq\alpha\cdot\delta$$

The first step will be to compare $\tilde{g}_\alpha(\phi)$ and $g_\alpha(f)$. Similarly to the proof of the lower bound of Theorem 3 from [2], we can define the function

$$h(z):=\int_{\mathcal{K}}\mathbf{1}_{f=z}.$$

Recall that h is non-zero only on a countable number of values which we will denote by z_i for $i\geq 1$. Then,

$$\begin{aligned}\left|\int_{\mathcal{K}}\tilde{g}_\alpha(\phi)-g_\alpha(f)\right|&=\int_{f\notin\{z_1,\cdots,z_Z\}}\left|\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha}-f(x)\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha}\right|\\&\leq\int_{f\notin\{z_1,\cdots,z_Z\}}|\phi(x)-f(x)|\left\|\sqrt{f}\right\|_1^{\alpha/2}+\phi(x)\cdot\varepsilon\alpha\max\left(\left\|\sqrt{f}\right\|_1^{\alpha-1},\left\|\sqrt{\phi}\right\|_1^{\alpha-1}\right)\\&\leq(1+\alpha)\delta.\end{aligned}$$

We now compare $g_\alpha(\phi)$ to $\phi(\cdot)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(\cdot)}\right)^{\alpha}$. For all $1\leq k\leq m^2$, we will denote $\eta_k:=\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi_k}$. We now use the fact that all ϕ_k are distinct. By definition of $g_\alpha(\phi)$,

$$\begin{aligned}\int_{\mathcal{K}}\left|g_\alpha(\phi,x)-\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right|dx&=\sum_{k=1}^{m^2}\phi_k\int_{t=0}^{\mathcal{A}(Q_k)}\left[(\eta_k+t)^\alpha-\eta_k^\alpha\right]dt\\&\leq\sum_{k=1}^{m^2}\frac{\phi_k}{m^2}\cdot\frac{\alpha}{m^2}\left(\int_{\mathcal{K}}\sqrt{\phi}\right)^{\alpha-1}\\&\leq\frac{\alpha}{m^2}.\end{aligned}$$

We take m sufficiently large so that the left term can be upper bounded by δ. We now turn to comparing $\tilde{g}_\alpha(\phi)$ and $\phi(\cdot)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(\cdot)}\right)^{\alpha}$.

$$\begin{aligned}&\left|\int_{\mathcal{K}}\left[\tilde{g}_\alpha(\phi,x)-\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right]dx\right|\\&\qquad\leq\sum_{i\geq 1}\left|\int_{f=z_i}\left[\tilde{g}_\alpha(\phi,x)-\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right]dx\right|\\&\qquad+\left|\int_{f\notin\{z_i,\,i\geq 1\}}\left[\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha}-\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right]dx\right|.\qquad(2.1)\end{aligned}$$

Let us analyze the second term in the right-hand side of the inequality.

$$\begin{aligned}
&\left|\int_{f\notin\{z_i,\, i\geq 1\}}\left[\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{f<f(x)}\right)^{\alpha}-\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right]dx\right| \\
&\qquad\leq \int_{f\notin\{z_i,\, i\geq 1\}}\phi(x)\cdot\alpha\left\|\sqrt{\phi}\right\|_1^{\alpha-1}\int_{\mathcal{K}}\sqrt{\phi}\cdot\left|\mathbf{1}_{f<f(x)}-\mathbf{1}_{\phi<\phi(x)}\right|dx \\
&\qquad\leq \alpha\iint_{\mathcal{K}^2}\phi(x)\sqrt{\phi(y)}\left|\mathbf{1}_{f(y)<f(x)}-\mathbf{1}_{\phi(y)<\phi(x)}\right|dxdy \\
&\qquad\leq \alpha\iint_{\mathcal{K}^2}\phi(x)\sqrt{\phi(y)}\mathbf{1}_{|f(y)-f(x)|\leq 3\varepsilon}\mathbf{1}_{f(x)\neq f(y)}dxdy.
\end{aligned}$$

By the dominated convergence theorem, this term vanishes as $\varepsilon\to 0$. We take $0<\varepsilon\leq\delta$ sufficiently small such that this term is upper bounded by δ. We now turn to the first term of Equation (2.1). For $1\leq i\leq Z$, denote by $\eta_i := \int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<z_i}$. Then,

$$\begin{aligned}
&\sum_i\left|\int_{f(x)=z_i}\left[\tilde{g}_\alpha(\phi,x)-\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right]dx\right| \\
&=\sum_i\left|\int_0^{\mathcal{A}(f=z_i)}z_i\left(\eta_i+\sqrt{z_i}t\right)^{\alpha}dt-\int_{f(x)=z_i}\phi(x)\left(\int_{\mathcal{K}}\sqrt{\phi}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right| \\
&\leq\sum_i z_i\left|\int_0^{h(z_i)}\left(\eta_i+\sqrt{z_i}t\right)^{\alpha}dt-\int_{f(x)=z_i}\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right|+\varepsilon\cdot\left\|\sqrt{f}\right\|_1^{\alpha} \\
&+\int_{\mathcal{K}}\phi(x)\cdot\alpha\max\left(\left\|\sqrt{\phi}\right\|_1^{\alpha-1},\left\|\sqrt{f}\right\|_1^{\alpha-1}\right)\varepsilon dx \\
&\leq\sum_i z_i\left|\int_0^{h(z_i)}\left(\eta_i+\sqrt{z_i}t\right)^{\alpha}dt-\int_{f(x)=z_i}\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right|+(1+\alpha)\delta.
\end{aligned}$$

For $x\in\mathcal{K}$ such that $f(x)=z_i$,

$$\begin{aligned}
\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)}&=\int_{|f-f(x)|>3\varepsilon}\sqrt{f}\cdot\mathbf{1}_{f<f(x)}+\int_{|f-f(x)|\leq 3\varepsilon}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)} \\
&=\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{f<z_i-3\varepsilon}+\sqrt{z_i}\int_{f=z_i}\mathbf{1}_{\phi<\phi(x)}+\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)}\mathbf{1}_{|f-z_i|\leq 3\varepsilon}\mathbf{1}_{f\neq z_i}.
\end{aligned}$$

Therefore,

$$\left|\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)}-\eta_i-\sqrt{z_i}\int_{f=z_i}\mathbf{1}_{\phi<\phi(x)}\right|\leq 2\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{|f-z_i|\leq 3\varepsilon}\mathbf{1}_{f\neq z_i}.$$

We can use this estimate for the following upper bound.

$$\sum_i z_i\left|\int_{f(x)=z_i}\left[\left(\eta_i+\sqrt{z_i}\int_{f=z_i}\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}-\left(\int_{\mathcal{K}}\sqrt{f}\cdot\mathbf{1}_{\phi<\phi(x)}\right)^{\alpha}\right]dx\right|$$

$$\leq 2\alpha \cdot \sum_i z_i \int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{|f-z_i|\leq 3\varepsilon} \mathbf{1}_{f\neq z_i}.$$

All terms in the right-hand side sum vanish as $\varepsilon \to 0$ by the dominated convergence theorem. Furthermore, the total sum is dominated by 2α. By monotone convergence, the sum vanishes as $\varepsilon \to 0$. Let us take $\varepsilon > 0$ sufficiently small such that the left-hand term is upper bounded by δ. The last term to analyze is

$$\sum_i z_i \int_{f(x)=z_i} \left(\eta_i + \sqrt{z_i} \int_{f=z_i} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha dx.$$

Let us take $Z \geq 1$ sufficiently large such that

$$\left| \sum_{i>Z} z_i h(z_i) \right| \leq \delta.$$

In particular, we can restrict the analysis to terms $1 \leq i \leq Z$ since

$$\sum_{i>Z} z_i \left| \int_0^{h(z_i)} (\eta_i + \sqrt{z_i} t)^\alpha \, dt - \int_{f(x)=z_i} \left(\int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{\phi<\phi(x)} \right)^\alpha \right| \leq 2 \sum_{i>Z} z_i h(z_i) \leq 2\delta.$$

Because $\{x : \; f(x) = z_i\}$ is measurable, for any tolerance $\varepsilon > 0$, there exists $m_0 \geq 1$ arbitrarily large and a set of sub-squares $E_i \subset \{1, \cdots, m^2\}$ such that

$$\left\| \mathbf{1}_{f=z_i} - \mathbf{1}_{\bigcup_{k\in E_i} Q_k} \right\|_1 \leq \frac{\delta}{Z}.$$

for all $1 \leq i \leq Z$. Then,

$$\begin{aligned}
&\sum_{i\leq Z} \left| \int_{f(x)=z_i} \left(\eta_i + \sqrt{z_i} \int_{f=z_i} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha dx - \int_{\bigcup_{k\in E_i} Q_k} \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha dx \right| \\
&\leq \sum_{i\leq Z} \int_{f(x)=z_i} \left| \left(\eta_i + \sqrt{z_i} \int_{f=z_i} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha - \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha \right| dx \\
&+ \sum_{i\leq Z} \int_{\mathcal{K}} \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha \left| \mathbf{1}_{f(x)=z_i} - \mathbf{1}_{\bigcup_{k\in E_i} Q_k} \right| dx \\
&\leq \sum_{i\leq Z} h(z_i) \cdot \sqrt{z_i} \varepsilon \cdot \alpha + \sum_{i\leq Z} \frac{\delta}{Z} \\
&\leq (1+\alpha)\delta.
\end{aligned}$$

Now note that because values of ϕ on each sub-square ϕ_k are all distinct, then

$$\int_{\bigcup_{k\in E_i} Q_k} \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi<\phi(x)} \right)^\alpha dx = \sum_{k=1}^{|E_i|} \frac{1}{m^2} \left(\eta_i + \sqrt{z_i} \frac{k-1}{m^2} \right)^\alpha.$$

Therefore,

$$\begin{aligned}
&\left| \int_0^{|E_i|/m^2} (\eta_i + \sqrt{z_i} t)^\alpha \, dt - \int_{\bigcup_{k\in E_i} Q_k} \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi < \phi(x)} \right)^\alpha dx \right| \\
&\le \sum_{k=1}^{|E_i|} \frac{1}{m^2} \left[\left(\eta_i + \sqrt{z_i} \frac{k}{m^2} \right)^\alpha - \left(\eta_i + \sqrt{z_i} \frac{k-1}{m^2} \right)^\alpha \right] \\
&\le \frac{|E_i|}{m^2} \frac{\alpha}{m^2} (\eta_i + \sqrt{z_i})^{\alpha-1} \\
&\le \left(h(z_i) + \frac{\delta}{Z} \right) \frac{\alpha}{m^2}.
\end{aligned}$$

We are now ready to merge all our estimates together.

$$\begin{aligned}
&\sum_{i\le Z} \left| \int_0^{h(z_i)} (\eta_i + \sqrt{z_i} t)^\alpha \, dt - \int_{f(x)=z_i} \left(\int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{\phi < \phi(x)} \right)^\alpha \right| \\
&\le \sum_{i\le Z} \frac{\delta}{Z} + \sum_{i\le Z} \left| \int_0^{|E_i|/m^2} (\eta_i + \sqrt{z_i} t)^\alpha \, dt - \int_{\bigcup_{k\in E_i} Q_k} \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi < \phi(x)} \right)^\alpha dx \right| \\
&+ \sum_{i\le Z} \left| \int_{\bigcup_{k\in E_i} Q_k} \left(\eta_i + \sqrt{z_i} \int_{\bigcup_{k\in E_i} Q_k} \mathbf{1}_{\phi < \phi(x)} \right)^\alpha dx - \int_{f(x)=z_i} \left(\eta_i + \sqrt{z_i} \int_{f=z_i} \mathbf{1}_{\phi < \phi(x)} \right)^\alpha dx \right| \\
&+ \sum_{i\le Z} \left| \int_{f(x)=z_i} \left[\left(\eta_i + \sqrt{z_i} \int_{f=z_i} \mathbf{1}_{\phi < \phi(x)} \right)^\alpha - \left(\int_{\mathcal{K}} \sqrt{f} \cdot \mathbf{1}_{\phi < \phi(x)} \right)^\alpha \right] dx \right| \\
&\le (3+\alpha)\delta + \frac{(1+\delta)\alpha}{m^2}.
\end{aligned}$$

We can then take m sufficiently large so that $\frac{(1+\delta)\alpha}{m^2} \le \delta$. Finally, going back to Eq 2.1,

$$\left| \int_{\mathcal{K}} \left[\tilde{g}_\alpha(\phi, x) - \phi(x) \left(\int_{\mathcal{K}} \sqrt{\phi} \cdot \mathbf{1}_{\phi < \phi(x)} \right)^\alpha \right] dx \right| \le (8 + 2\alpha)\delta$$

We now conclude by noting that

$$\begin{aligned}
\left| \int_{\mathcal{K}} g_\alpha(\phi) - g_\alpha(f) \right| &\le \left| \int_{\mathcal{K}} \left[g_\alpha(\phi, x) - \phi(x) \left(\int_{\mathcal{K}} \sqrt{\phi} \cdot \mathbf{1}_{\phi < \phi(x)} \right)^\alpha \right] dx \right| \\
&+ \left| \int_{\mathcal{K}} \left[\phi(x) \left(\int_{\mathcal{K}} \sqrt{\phi} \cdot \mathbf{1}_{\phi < \phi(x)} \right)^\alpha - \tilde{g}_\alpha(\phi, x) \right] dx \right| \\
&+ \left| \int_{\mathcal{K}} \tilde{g}_\alpha(\phi) - g_\alpha(f) \right| \\
&\le (10 + 3\alpha)\delta
\end{aligned}$$

This is true for any $\delta > 0$. This ends the proof of the lemma. $\square$

References

[1] Bertsimas, Dimitris ; Farias, Vivek ; Trichakis, Nikolaos: The price of fairness. In: *Operations research* 59 (2011), Nr. 1, S. 17–31

[2] Blanchard, Moïse ; Jacquillat, Alexandre ; Jaillet, Patrick: Probabilistic bounds on the k-Traveling Salesman Problem and the Traveling Repairman Problem. (2022)