Recent Advances on Approximation Algorithms in Doubling Metrics

Shaofeng Jiang

The University of Hong Kong
PTASes in Euclidean Spaces
Motivation: Approximation Algorithms in Euclidean Spaces

PTASes in Euclidean Spaces

- PTAS for Steiner tree, TSP [Aro96].
Motivation: Approximation Algorithms in Euclidean Spaces

PTASes in Euclidean Spaces

- PTAS for Steiner tree, TSP [Aro96].
- PTAS for $k$-median [ARR98].
Motivation: Approximation Algorithms in Euclidean Spaces

PTASes in Euclidean Spaces

- PTAS for Steiner tree, TSP [Aro96].
- PTAS for $k$-median [ARR98].
- PTAS for Steiner forest [BKM08].
Motivation: PTASes in Euclidean Spaces

Key Technique: Randomized Quad-tree Subdivision [Aro96]

In addition to the bounded dimensionality, it uses geometry and vector representation.

Figure: Quad-tree subdivision with random shifts in x and y directions.
Motivation: PTASes in Euclidean Spaces

Key Technique: Randomized Quad-tree Subdivision [Aro96]
In addition to the bounded dimensionality, it uses geometry and vector representation.

Figure: Quad-tree subdivision with random shifts in x and y directions.

Question: PTAS using the bounded dimensionality only?
Doubling Dimension: Dimensionality of General Metrics

**Definition (doubling dimension)**

A metric space has doubling dimension at most $k$, if any ball can be covered by at most $2^k$ balls of half the radius.
Definition (doubling dimension)

A metric space has doubling dimension at most $k$, if any ball can be covered by at most $2^k$ balls of half the radius.
Doubling Dimension: Dimensionality of General Metrics

Definition (doubling dimension)

A metric space has doubling dimension at most \( k \), if any ball can be covered by at most \( 2^k \) balls of half the radius.

Some Properties [Ass83]

- \( \mathbb{R}^k \) equipped with \( \ell_p \) has doubling dimension \( \Theta(k) \).
- A subset of \( \mathbb{R}^k \) equipped with \( \ell_p \) has doubling dimension \( O(k) \).
- A set of \( n \) points has doubling dimension \( \log n \).
Outline

1. Traveling Salesman Problem (TSP) in doubling metrics.
Outline

1. Traveling Salesman Problem (TSP) in doubling metrics.
   ▶ A QPTAS [Tal04].

2. Steiner Forest Problem (SFP) in doubling metrics.
   ▶ A PTAS [CHJ16].

3. Open questions.
1. Traveling Salesman Problem (TSP) in doubling metrics.
   - A QPTAS [Tal04].
   - A PTAS [BGK12] and its generalization [CHJ16].
Outline

1. Traveling Salesman Problem (TSP) in doubling metrics.
   ▶ A QPTAS [Tal04].
   ▶ A PTAS [BGK12] and its generalization [CHJ16].
2. Steiner Forest Problem (SFP) in doubling metrics.
   ▶ A PTAS [CHJ16].
Outline

1. Traveling Salesman Problem (TSP) in doubling metrics.
   - A QPTAS [Tal04].
   - A PTAS [BGK12] and its generalization [CHJ16].

2. Steiner Forest Problem (SFP) in doubling metrics.
   - A PTAS [CHJ16].

3. Open questions.
Traveling Salesman Problem

Given: metric $M := (X, d)$ with doubling dimension $k$, and $V \subset X$. Goal: find a minimum length tour that visits every point in $V$. 
Traveling Salesman Problem

Given: metric $M := (X, d)$ with doubling dimension $k$, and $V \subset X$. Goal: find a minimum length tour that visits every point in $V$.

Figure: An instance.
Traveling Salesman Problem

Given: metric $M := (X, d)$ with doubling dimension $k$, and $V \subseteq X$. Goal: find a minimum length tour that visits every point in $V$.

Figure: An instance and its solution.
Talwar [Tal04] gave a QPTAS for the TSP in doubling metrics.
Talwar [Tal04] gave a QPTAS for the TSP in doubling metrics.

**Theorem (QPTAS [Tal04])**

There is a randomized algorithm that returns a $(1 + \epsilon)$-approximate solution for the TSP with constant probability, running in time

$$\text{poly}(n) \cdot \left( \frac{k \log n}{\epsilon} \right)^{\frac{k \log n}{\epsilon}^k},$$

where $k$ is the doubling dimension, and $n$ is the number of points in the TSP instance.
Basic Structure: Packing, Covering and Nets

Definition (packing)

$S \subset X$ is a $\rho$-packing, if the distance between any two different points in $S$ is at least $\rho$. 
Basic Structure: Packing, Covering and Nets

Definition (packing)

$S \subset X$ is a $\rho$-packing, if the distance between any two different points in $S$ is at least $\rho$.

Fact (packing property)

*If $S$ is a $\rho$-packing with doubling dimension $k$ and diameter at most $D$, then $|S| \leq \left(\frac{2D}{\rho}\right)^k$.*
Basic Structure: Packing, Covering and Nets

Definition (packing)

$S \subset X$ is a $\rho$-packing, if the distance between any two different points in $S$ is at least $\rho$.

Definition (covering)

$S \subset X$ is a $\rho$-covering if $\forall u \in X, \exists v \in S : d(u, v) \leq \rho$. 
Basic Structure: Packing, Covering and Nets

Definition (packing)
$S \subset X$ is a $\rho$-packing, if the distance between any two different points in $S$ is at least $\rho$.

Definition (covering)
$S \subset X$ is a $\rho$-covering if $\forall u \in X, \exists v \in S : d(u, v) \leq \rho$.

Definition (net)
$S \subset X$ is a $\rho$-net if $S$ is both a $\rho$-covering and $\rho$-packing.
Basic Structure: Hierarchical Nets

Geometric Distance Scales
Let $s \geq 2$ be the scaling factor. We say scale $s^i$ is of height $i$.

Hierarchical Nets
A greedy algorithm can construct

$$N_L \subset N_{L-1} \subset \ldots \subset N_1 \subset N_0 = X$$

such that $N_i$ is a $s^i$-net for $X$. 
Basic Structure: Hierarchical Nets

Geometric Distance Scales
Let \( s \geq 2 \) be the scaling factor. We say scale \( s^i \) is of height \( i \).

Hierarchical Nets
A greedy algorithm can construct
\[
N_L \subset N_{L-1} \subset \ldots \subset N_1 \subset N_0 = X
\]
such that \( N_i \) is a \( s^i \)-net for \( X \).

Bounded Instances
W.l.o.g., one can assume the minimum intra-point distance is 1, and the diameter of the metric is \( \text{poly}(n) \).
\[
L := \log_s(\text{poly}(n)) = O(\log n)
\]
is sufficient.
Randomized Hierarchical Decomposition [ABN06]

Single Scale Decomposition

Given: a height $i$, a permutation of points $\tau$, and $N_i$.
Return: a random partition $\Pi_i$ of $X$. 

1. Each point $u \in N_i$ corresponds to a part (which we call cluster) in $\Pi_i$.
2. Sample random radii $h(i)_u \in [0, s_i]$ from a truncated exponential distribution.
3. Define the ball $B(i)_u := \{v \in X: d(u, v) \leq s_i + h(i)_u\}$.
4. A point $p \in X$ belongs to the cluster centered at $u$, if $B(i)_u$ is the first w.r.t $\tau$ that contains $p$. 

Randomized Hierarchical Decomposition [ABN06]

Single Scale Decomposition

Given: a height $i$, a permutation of points $\tau$, and $N_i$.
Return: a random partition $\Pi_i$ of $X$.

1. Each point $u \in N_i$ corresponds to a part (which we call a cluster) in $\Pi_i$. 
Randomized Hierarchical Decomposition [ABN06]

Single Scale Decomposition

Given: a height $i$, a permutation of points $\tau$, and $N_i$.
Return: a random partition $\Pi_i$ of $X$.

1. Each point $u \in N_i$ corresponds to a part (which we call cluster) in $\Pi_i$.

2. Sample random radii $h_u^{(i)} \in [0, s^i]$ from a truncated exponential distribution.
Randomized Hierarchical Decomposition [ABN06]

**Single Scale Decomposition**

Given: a height $i$, a permutation of points $\tau$, and $N_i$.
Return: a random partition $\Pi_i$ of $X$.

1. Each point $u \in N_i$ corresponds to a part (which we call cluster) in $\Pi_i$.
2. Sample random radii $h_u^{(i)} \in [0, s^i]$ from a truncated exponential distribution.
3. Define the ball $B_u^{(i)} := \{ v \in X : d(u, v) \leq s^i + h_u^{(i)} \}$. 


Randomized Hierarchical Decomposition [ABN06]

Single Scale Decomposition

Given: a height $i$, a permutation of points $\tau$, and $N_i$.
Return: a random partition $\Pi_i$ of $X$.

1. Each point $u \in N_i$ corresponds to a part (which we call cluster) in $\Pi_i$.

2. Sample random radii $h_u^{(i)} \in [0, s^i]$ from a truncated exponential distribution.

3. Define the ball $B_u^{(i)} := \{ v \in X : d(u, v) \leq s^i + h_u^{(i)} \}$.

4. A point $p \in X$ belongs to the cluster centered at $u$, if $B_u^{(i)}$ is the first w.r.t $\tau$ that contains $p$. 
An Example

Figure: $N_i$ is marked in red.
Figure: First ball.
An Example

Figure: Second ball.
An Example

Figure: Third ball: some points are included in previous balls
Hierarchical Decomposition [ABN06]

1. For each height $i = 0, 1, 2, \ldots, L$, construct $\{\Pi_i\}_{i=0}^L$.

2. Base Case. At height $L$, define height-$L$ clusters as parts in $\Pi_L$.

3. General Case. For $i = L-1, L-2, \ldots, 0$, form height-$i$ clusters by partitioning height-$(i+1)$ clusters according to $\Pi_i$.

Note: Each height-$i$ cluster is determined by all random radius for $u \in N_j$ and $j \geq i$.

Theorem: For $S \subset X$, $\Pr[S$ is cut at height $i] \leq O(k) \cdot \text{Diam}(S)_{si}$. 
Hierarchical Decomposition [ABN06]

1. For each height $i = 0, 1, 2, \ldots, L$, construct $\{\Pi_i\}_i$. 
Hierarchical Decomposition [ABN06]

1. For each height $i = 0, 1, 2, \ldots, L$, construct $\{\Pi_i\}_i$.

2. **Base Case.** At height $L$, define height-$L$ clusters as parts in $\Pi_L$.
Hierarchical Decomposition [ABN06]

1. For each height $i = 0, 1, 2, \ldots, L$, construct $\{\Pi_i\}_i$.
2. **Base Case.** At height $L$, define height-$L$ clusters as parts in $\Pi_L$.
3. **General Case.** For $i = L - 1, L - 2, \ldots, 0$, form height-$i$ clusters by partitioning height-$(i + 1)$ clusters according to $\Pi_i$. 

Note: Each height-$i$ cluster is determined by all random radius for $u \in N_j$ and $j \geq i$.

Theorem: For $S \subset X$, $\Pr[\text{S is cut at height } i] \leq O(k) \cdot \text{Diam}(S)$. 


Hierarchical Decomposition [ABN06]

1. For each height $i = 0, 1, 2, \ldots, L$, construct $\{\Pi_i\}_i$.

2. **Base Case.** At height $L$, define height-$L$ clusters as parts in $\Pi_L$.

3. **General Case.** For $i = L - 1, L - 2, \ldots, 0$, form height-$i$ clusters by partitioning height-$(i + 1)$ clusters according to $\Pi_i$.

**Note**
Each height-$i$ cluster is determined by all random radius for $u \in N_j$ and $j \geq i$. 
Hierarchical Decomposition [ABN06]

1. For each height \( i = 0, 1, 2, \ldots, L \), construct \( \{\Pi_i\}_i \).

2. **Base Case.** At height \( L \), define height-\( L \) clusters as parts in \( \Pi_L \).

3. **General Case.** For \( i = L - 1, L - 2, \ldots, 0 \), form height-\( i \) clusters by partitioning height-(\( i + 1 \)) clusters according to \( \Pi_i \).

**Note**
Each height-\( i \) cluster is determined by all random radius for \( u \in N_j \) and \( j \geq i \).

**Theorem**
For \( S \subset X \), \( \Pr[S \text{ is cut at height } i] \leq O(k) \cdot \frac{Diam(S)}{s^i} \).
QPTAS: General Strategy

- Prove a structural property that there exists a "simple structured" solution that is $(1 + \epsilon)$-approximate.
- Use a dynamic program to find the optimal "simple structured" solution.
QPTAS: General Strategy

- Prove a structural property that there exists a “simple structured” solution that is $(1 + \epsilon)$-approximate.
QPTAS: General Strategy

- Prove a structural property that there exists a “simple structured” solution that is $(1 + \epsilon)$-approximate.
- Use a dynamic program to find the optimal “simple structured” solution.
Portal Respecting Solutions [Tal04]

Fix a hierarchical decomposition.
Portal Respecting Solutions [Tal04]

Fix a hierarchical decomposition.
For each cluster $C$ of height $i$, define points $N_j$ that cover $C$ as portals, for $j < i$ to be picked later.
Portal Respecting Solutions [Tal04]

Fix a hierarchical decomposition.
For each cluster $C$ of height $i$, define points $N_j$ that cover $C$ as portals, for $j < i$ to be picked later.

Definition (Portal Respecting Solution)

A solution is *portal respecting*, if it crosses each cluster via portals only.
Definition ($(m, r)$-light solution)

A solution $F$ is $(m, r)$-light, if

- $F$ is portal respecting with respect to at most $m$ predefined portals for each cluster;
- $F$ crosses each cluster via at most $r$ portals which we call active portals.

Example

A partial $(10, 4)$-light solution.
Theorem (Structural Property)

For each hierarchical decomposition $\Pi$, there is a $(m, r)$-light solution $F$, such that $w(F) \leq (1 + \epsilon) \cdot OPT$ with constant probability, where

$$m := (O\left(\frac{skL}{\epsilon}\right))^k, \quad r := (O\left(\frac{skL}{\epsilon}\right))^k.$$
Theorem (Structural Property)

For each hierarchical decomposition $\Pi$, there is a $(m, r)$-light solution $F$, such that $w(F) \leq (1 + \epsilon) \cdot OPT$ with constant probability, where

$$m := (O\left(\frac{skL}{\epsilon}\right))^k, \quad r := (O\left(\frac{skL}{\epsilon}\right))^k.$$

Note

Both $m$ and $r$ are polylog($n$).
Step 1: Portal Respecting

Portals for a height-\(i\) cluster \(C\):
the subset of \(N_j\) that covers \(C\), for \(s^j < \Theta\left( \frac{\epsilon}{kL} \right) \cdot s^i \leq s^{j+1}\).
So \(m = (O\left( \frac{skL}{\epsilon} \right))^k\).
Step 1: Portal Respecting

Portals for a height-\(i\) cluster \(C\):
the subset of \(N_j\) that covers \(C\), for \(s^j < \Theta\left(\frac{\epsilon}{kL}\right) \cdot s^i \leq s^{i+1}\).
So \(m = (O\left(\frac{skL}{\epsilon}\right))^k\).

Consider an edge \((u, v)\) in OPT.
Let \(i\) be the largest height that \((u, v)\) is cut.

**Rerouting**

Define \(u', v'\) be the nearest from \(u, v\) in \(N_j\). Replace \((u, v)\) with \((u, u')\), \((u', v')\) and \((v, v')\). Do this for \((u, u')\) and \((v, v')\) recursively.
Step 1: Portal Respecting

Portals for a height-$i$ cluster $C$:
the subset of $N_j$ that covers $C$, for $s^j < \Theta\left(\frac{\epsilon}{kL}\right) \cdot s^i \leq s^{i+1}$.
So $m = \left(O\left(\frac{skL}{\epsilon}\right)\right)^k$.

Cost

\begin{align*}
d(u', v') & \leq d(u, v) + d(u, u') + d(v, v') \\
                     & \leq d(u, v) + 2s^j,
\end{align*}

and the recursive cost: $O(s^j)$.
Hence, the cost for $(u, v)$ first cut at height $i$ is at most $O(s^j) = O\left(\frac{\epsilon}{kL}\right) \cdot s^i$.
In expectation, this is at most $O(k) \cdot \frac{d(u, v)}{s^i} \cdot O\left(\frac{\epsilon}{kL}\right) \cdot s^i = O\left(\frac{\epsilon}{L}\right) \cdot d(u, v)$. 

Figure: Rerouting
Step 1: Portal Respecting

Portals for a height-$i$ cluster $C$: the subset of $N_j$ that covers $C$, for $s^j < \Theta\left(\frac{\epsilon}{kL}\right) \cdot s^i \leq s^{i+1}$.

So $m = (O\left(\frac{skL}{\epsilon}\right))^k$.

Union Bound

Since the largest height $i$ that $(u, v)$ is cut is random, a union bound is needed. Hence the expected cost is at most

$$\sum_{i=1}^{L} O\left(\frac{\epsilon}{L}\right) \cdot d(u, v) = O(\epsilon) \cdot d(u, v).$$
Step 2: Reducing Number of Crossings

Lemma (Patching)

Suppose $C$ is a cluster of height $i$ and the portal set is $R$. Any portal respecting tour $T$ can be modified to a portal respecting tour $T'$ such that

1. $T'$ visits all points that $T$ visits;
2. $T'$ crosses $C$ at most twice;
3. Comparing with $T$, the number of crossings for clusters of height at least $i$ does not increase in $T'$;
4. $w(T') \leq w(T) + O(w(MST(R)))$. 


Step 2: Reducing Number of Crossings

Lemma (Patching)

Suppose \( C \) is a cluster of height \( i \) and the portal set is \( R \). Any portal respecting tour \( T \) can be modified to a portal respecting tour \( T' \) such that

- \( T' \) visits all points that \( T \) visits;

- \( T' \) crosses \( C \) at most twice;

- Comparing with \( T \), the number of crossings for clusters of height at least \( i \) does not increase in \( T' \);

- \( w(T') \leq w(T) + O(w(MST(R))) \).
Step 2: Reducing Number of Crossings

Lemma (Patching)

Suppose $C$ is a cluster of height $i$ and the portal set is $R$. Any portal respecting tour $T$ can be modified to a portal respecting tour $T'$ such that

- $T'$ visits all points that $T$ visits;
- $T'$ crosses $C$ at most twice;
Lemma (Patching)

Suppose $C$ is a cluster of height $i$ and the portal set is $R$. Any portal respecting tour $T$ can be modified to a portal respecting tour $T'$ such that

- $T'$ visits all points that $T$ visits;
- $T'$ crosses $C$ at most twice;
- Comparing with $T$, the number of crossings for clusters of height at least $i$ does not increase in $T'$;
Step 2: Reducing Number of Crossings

Lemma (Patching)

Suppose \( C \) is a cluster of height \( i \) and the portal set is \( R \). Any portal respecting tour \( T \) can be modified to a portal respecting tour \( T' \) such that

- \( T' \) visits all points that \( T \) visits;
- \( T' \) crosses \( C \) at most twice;
- Comparing with \( T \), the number of crossings for clusters of height at least \( i \) does not increase in \( T' \);
- \( w(T') \leq w(T) + O(w(MST(R))) \).
Step 2: Reducing Number of Crossings

Modification
In the order of $i = L, L - 1, \ldots, 1, 0$, for each height-$i$ cluster $C$, if $C$ has more than $r$ crossings, then we apply the Patching Lemma.
Step 2: Reducing Number of Crossings

Modification
In the order of $i = L, L - 1, \ldots, 1, 0$, for each height-$i$ cluster $C$, if $C$ has more than $r$ crossings, then we apply the Patching Lemma.

Lemma (Small MST)
Suppose $S \subset X$ is of diameter $l$. Then $w(MST(S)) \leq 4l \cdot |S|^{1 - \frac{1}{k}}$. 
Step 2: Reducing Number of Crossings

Modification
In the order of $i = L, L - 1, \ldots, 1, 0$, for each height-$i$ cluster $C$, if $C$ has more than $r$ crossings, then we apply the Patching Lemma.

Lemma (Small MST)
Suppose $S \subset X$ is of diameter $l$. Then $w(MST(S)) \leq 4l \cdot |S|^{1 - \frac{1}{k}}$.

Cost
By the small MST Lemma, the modification costs at most

$$O(s^i) \cdot r'^{1 - \frac{1}{k}},$$

for a cluster $C$ at height $i$ with $r'$ crossings.
Charging Argument

Charging Scheme

We charge this to each of the crossing edges, so each edge takes 

\[ O(s^i) \cdot \frac{r'^{1 - \frac{1}{k}}}{r'} = O(s^i) \cdot r'^{-\frac{1}{k}} \leq O(s^i) \cdot r^{-\frac{1}{k}}. \]
Charging Argument

Charging Scheme
We charge this to each of the crossing edges, so each edge takes
\[ O(s^i) \cdot \frac{r'^{1-\frac{1}{k}}}{r'} = O(s^i) \cdot r'^{-\frac{1}{k}} \leq O(s^i) \cdot r^{-\frac{1}{k}}. \]

Accounting
Since an edge can be cut at most twice at any fixed height, the charge that an edge \((u, v)\) takes at height \(i\) is \(O(s^i) \cdot r^{-\frac{1}{k}}\).
Charging Argument

Charging Scheme
We charge this to each of the crossing edges, so each edge takes
\[ O(s^i) \cdot \frac{r'^{1-\frac{1}{k}}}{r'} = O(s^i) \cdot r'^{-\frac{1}{k}} \leq O(s^i) \cdot r^{-\frac{1}{k}}. \]

Accounting
Since an edge can be cut at most twice at any fixed height, the
charge that an edge \((u, v)\) takes at height \(i\) is \(O(s^i) \cdot r^{-\frac{1}{k}}\).
So, the expected cost is at most
\[
\sum_{i=1}^{L} O(k) \cdot \frac{d(u, v)}{s^i} \cdot s^i \cdot r^{-\frac{1}{k}} = O(kL) \cdot r^{-\frac{1}{k}} \cdot d(u, v)
\]
\[
\leq O(\epsilon) \cdot d(u, v),
\]
recalling \(r := (O(\frac{skL}{\epsilon}))^k\).
A Dynamic Program

A DP is applied to find the optimal \((m, r)\)-light solution, for any \(\Pi\).
A Dynamic Program

A DP is applied to find the optimal \((m, r)\)-light solution, for any \(\Pi\).
Subproblem: \((C, R, P)\).

- \(C\) denotes a cluster.
- \(R\) denotes the active portals.
- \(P\) is a collection of pairs of \(R\): pair \((u, v)\) \(\in P\) means a portion of the tour enters and leaves \(C\) at \(u\) and \(v\).
A Dynamic Program

A DP is applied to find the optimal \((m, r)\)-light solution, for any \(\Pi\).
Subproblem: \((C, R, P)\).

- \(C\) denotes a cluster.
- \(R\) denotes the active portals.
- \(P\) is a collection of pairs of \(R\): pair \((u, v)\) \(\in P\) means a portion of the tour enters and leaves \(C\) at \(u\) and \(v\).

Example

\(R = \{A, B, D\}\). \(P = \{(A, B), (D, D)\}\).
A Dynamic Program

A DP is applied to find the optimal \((m, r)\)-light solution, for any \(\Pi\).
Subproblem: \((C, R, P)\).

- \(C\) denotes a cluster.
- \(R\) denotes the active portals.
- \(P\) is a collection of pairs of \(R\): pair \((u, v) \in P\) means a portion of the tour enters and leaves \(C\) at \(u\) and \(v\).

Running Time Analysis

Number of \((R, P)\):
\[
\binom{m}{\leq r} \cdot r^r \approx (mr)^r.
\]
Plugging in \(m = \text{poly log } n, r = \text{poly log } n\),
\[
(mr)^r \approx (\log n)^{\text{poly log } n}.
\]
A ground breaking result by Bartal, Gottlieb and Krauthgamer [BGK12] gives a PTAS for the TSP in doubling metrics.

Running time:

\[ n^{O(1)^k} \cdot \exp\left(O\left(\frac{1}{\epsilon}\right)^{k^2} \cdot \sqrt{\log n}\right). \]
Our Work: TSPN

Our work [CJ16] gives a PTAS for (a special version of) TSP with neighborhoods (TSPN) in doubling metrics.
Our Work: TSPN

Our work [CJ16] gives a PTAS for (a special version of) TSP with neighborhoods (TSPN) in doubling metrics.

**TSPN**
Given: a collection of subsets of points (which are called regions).
Goal: find a lightest tour visiting at least one point in each regions.
Our Work: TSPN

Our work [CJ16] gives a PTAS for (a special version of) TSP with neighborhoods (TSPN) in doubling metrics.

TSPN
Given: a collection of subsets of points (which are called regions).
Goal: find a lightest tour visiting at least one point in each regions.
Before this work, only a QPTAS [CE10] is known for the problem.
Improving the Running Time

The running time for the TSP is improved to

$$n^{O(1)^k} \cdot \exp(O\left(\frac{k}{\epsilon}\right)^k \cdot \sqrt{\log n}),$$

compared with [BGK12]

$$n^{O(1)^k} \cdot \exp(O\left(\frac{1}{\epsilon}\right)^{k^2} \cdot \sqrt{\log n}).$$
Technical Contribution

Improving the Running Time
The running time for the TSP is improved to

\[ n^{O(1)^k \cdot \exp(O(\frac{k}{\epsilon})^k \cdot \sqrt{\log n})}, \]

compared with [BGK12]

\[ n^{O(1)^k \cdot \exp(O(\frac{1}{\epsilon})^{k^2} \cdot \sqrt{\log n})}. \]

Generalized Framework
Our framework applies to TSP, TSPN, and also the Steiner Forest Problem (which we shall see later).
Key Notion: Sparsity [BGK12]

Definition
A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u,3s^i)}) \leq q \cdot s^i$, where $F|_S$ for some $S$ is the subgraph of $F$ induced by vertices in $S$.

(a) Sparse

(b) Less Sparse
An instance $q$-sparse if the instance has a (near) optimal solution that is $q$-sparse. Let $q_0 := \Theta\left(\frac{sk}{\epsilon}\right)\Theta(k)$. 

Framework [BGK12, CJ16]
Framework [BGK12, CJ16]

An instance $q$-sparse if the instance has a (near) optimal solution that is $q$-sparse. Let $q_0 := \Theta\left(\frac{sk}{\epsilon}\right)^{\Theta(k)}$.

1. A Reduction to Sparse Instances

If there is a PTAS for $q_0$-sparse instances, then there is a (randomized) PTAS for general instances.
Framework [BGK12, CJ16]

An instance $q$-sparse if the instance has a (near) optimal solution that is $q$-sparse. Let $q_0 := \Theta\left(\frac{sk}{\epsilon}\right)\Theta(k)$.

1. A Reduction to Sparse Instances
If there is a PTAS for $q_0$-sparse instances, then there is a (randomized) PTAS for general instances.

2. A PTAS for Sparse Instances
There is a (randomized) PTAS for $q_0$-sparse instances.
PTAS for Sparse Instances: How Can Sparsity Help?

As in the QPTAS, consider the \((m, r)\)-light solutions.
As in the QPTAS, consider the \((m, r)\)-light solutions.

**Improved Structural Property**

There exists an \((m, r)\)-light solution \(F\), where

\[
m := (O\left(\frac{skL}{\epsilon}\right))^k, \quad r := O(q) \cdot 2^{O(k)} + (O\left(\frac{sk}{\epsilon}\right))^k,
\]

such that \(w(F) \leq (1 + \epsilon) \cdot \text{OPT}\).
As in the QPTAS, consider the \((m, r)\)-light solutions.

**Improved Structural Property**

There exists an \((m, r)\)-light solution \(F\), where

\[
m := (O\left(\frac{skL}{\epsilon}\right))^k, \quad r := O(q) \cdot 2^{O(k)} + (O\left(\frac{sk}{\epsilon}\right))^k,
\]

such that \(w(F) \leq (1 + \epsilon) \cdot \text{OPT}\).

**Note**

- Choosing parameters property, \(m = \text{polylog}(n)\), \(r = O(\log c n)\), where \(0 < c < 1\) is a small universal constant (say 0.00001).
- Recall that the number of subproblems is dominated by \((mr)^r\), but now \((mr)^r = (\log n)^{O(\log c n)} = 2^{O(\log c n + o(1)n)}\).
- Constant probability is actually required.
PTAS for Sparse Instances: How Can Sparsity Help?

As in the QPTAS, consider the \((m, r)\)-light solutions.

**Improved Structural Property**
There exists an \((m, r)\)-light solution \(F\), where

\[
m = (O(\frac{skL}{\epsilon}))^k, \quad r = O(q) \cdot 2^{O(k)} + (O(\frac{sk}{\epsilon}))^k,
\]

such that \(w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Note**
- Choosing parameters property, \(m = \text{polylog} n, \ r = O(\log^c n)\),
  where \(0 < c < 1\) is a small universal constant (say 0.00001).
PTAS for Sparse Instances: How Can Sparsity Help?

As in the QPTAS, consider the \((m, r)\)-light solutions.

**Improved Structural Property**
There exists an \((m, r)\)-light solution \(F\), where

\[
m := \left( O\left( \frac{skL}{\epsilon} \right) \right)^k, \quad r := O(q) \cdot 2^{O(k)} + \left( O\left( \frac{sk}{\epsilon} \right) \right)^k,
\]

such that \(w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Note**

- Choosing parameters property, \(m = \text{polylog} \ n\), \(r = O(\log^c n)\), where \(0 < c < 1\) is a small universal constant (say 0.00001).
- Recall that the number of subproblems is dominated by \((mr)^r\), but now \((mr)^r = (\log n)^{O(\log^c n)} = 2^{O(\log^{c+o(1)} n)}\).
As in the QPTAS, consider the \((m, r)\)-light solutions.

**Improved Structural Property**
There exists an \((m, r)\)-light solution \(F\), where

\[
m := \left( O\left( \frac{skL}{\epsilon} \right) \right)^k, \quad r := O(q) \cdot 2^{O(k)} + \left( O\left( \frac{sk}{\epsilon} \right) \right)^k,
\]

such that \(w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Note**

- Choosing parameters property, \(m = \text{polylog} n\), \(r = O(\log^c n)\), where \(0 < c < 1\) is a small universal constant (say 0.00001).
- Recall that the number of subproblems is dominated by \((mr)^r\), but now \((mr)^r = (\log n)^{O(\log^c n)} = 2^{O(\log^{c+o(1)} n)}\).
- Constant probability is actually required.
Portal Respecting

Apply the same procedure as in the QPTAS, so \( m := (O\left(\frac{skL}{\varepsilon}\right))^k \) suffices.
Proof Strategy

Portal Respecting

Apply the same procedure as in the QPTAS, so $m := (O(\frac{sL}{\varepsilon}))^k$ suffices.

Reducing Number of Crossings
Proof Strategy

Portal Respecting
Apply the same procedure as in the QPTAS, so $m := (O(\frac{skL}{\varepsilon}))^k$ suffices.

Reducing Number of Crossings

- For “long” edges, consider the net-respecting solution, and show that very few long edges can cross.
Proof Strategy

Portal Respecting
Apply the same procedure as in the QPTAS, so \( m := (O\left(\frac{skL}{\epsilon}\right))^k \) suffices.

Reducing Number of Crossings

- For “long” edges, consider the net-respecting solution, and show that very few long edges can cross.
- For “short” edges, use sparsity and give an improved charging argument.
Proof Strategy

Portal Respecting
Apply the same procedure as in the QPTAS, so \( m := (O(\frac{skL}{\epsilon}))^k \) suffices.

Reducing Number of Crossings
- For “long” edges, consider the net-respecting solution, and show that very few long edges can cross.
- For “short” edges, use sparsity and give an improved charging argument.

Definition (Long and Short Edges)
For height-\( i \) cluster \( C \), long edges \( > s^i \), short edges \( \leq s^i \).
Reducing Crossings: Long Edges

Definition (Net-respecting)

A graph $F$ is net-respecting, if for each edge $(u, v)$ of $F$, $u \in N_j$ and $v \in N_j$ for $s^j < \epsilon \cdot d(u, v) \leq s^{j+1}$.
Reducing Crossings: Long Edges

Definition (Net-respecting)
A graph $F$ is net-respecting, if for each edge $(u, v)$ of $F$, $u \in N_j$ and $v \in N_j$ for $s^j < \epsilon \cdot d(u, v) \leq s^{j+1}$.

Lemma (Net-respecting is w.l.o.g.)
For any graph $F$, there is a net-respecting graph $F'$ that visits all points visited by $F$ and $w(F') \leq (1 + O(\epsilon)) \cdot w(F)$. 
Reducing Crossings: Long Edges

Definition (Net-respecting)
A graph $F$ is net-respecting, if for each edge $(u, v)$ of $F$, $u \in N_j$ and $v \in N_j$ for $s_j < \epsilon \cdot d(u, v) \leq s_{j+1}$.

Lemma (Net-respecting is w.l.o.g.)

For any graph $F$, there is a net-respecting graph $F'$ that visits all points visited by $F$ and $w(F') \leq (1 + O(\epsilon)) \cdot w(F)$.

Claim (Few long edges can cross)

For each cluster, there are at most $(O(\frac{s}{\epsilon}))^k$ long crossing edges in a net-respecting solution.
Reducing Crossings: Long Edges

Definition (Net-respecting)
A graph $F$ is net-respecting, if for each edge $(u, v)$ of $F$, $u \in N_j$ and $v \in N_j$ for $s^j < \epsilon \cdot d(u, v) \leq s^{j+1}$.

Lemma (Net-respecting is w.l.o.g.)
For any graph $F$, there is a net-respecting graph $F'$ that visits all points visited by $F$ and $w(F') \leq (1 + O(\epsilon)) \cdot w(F)$.

Claim (Few long edges can cross)
For each cluster, there are at most $(O(\frac{s}{\epsilon}))^k$ long crossing edges in a net-respecting solution.

Proof.
Suppose $(u, v)$ is a long crossing edge of a height-$i$ cluster $C$. Then $u \in N_j$ and $v \in N_j$ for $s^j < \epsilon d(u, v) \leq s^{j+1}$.
Since $\text{Diam}(C) \leq O(s^i)$, $|N_j \cap C| \leq \left(\frac{O(s^i)}{s^j}\right)^k \leq (O(\frac{s}{\epsilon}))^k$. \qed
Short Edges: Using Sparsity

Definition (Recall: sparsity)

A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u,3s^i)}) \leq q \cdot s^i$. 

Observation

Consider $u \in N_i$ for some $i$, and a ball $B := B(u,s^i + h)$ where $h$ is sampled from $[0,s^i]$ uniformly at random. Define $X(h)$ the number of short edges cut by $B$. Then

$$\int_0^{s^i} X(h) \, dh \leq w(F|_{B(u,3s^i)}) \leq q \cdot s^i.$$

So $\Pr[X(h) \leq q^2] \geq 1/2$.

Good Radius

Sparsity implies that for each $u \in N_i$, $B(u,s^i + h)$ cuts at most $O(q)$ short edges with constant probability. In the following, we condition on such event for all $i$ and $u \in N_i$. 


Short Edges: Using Sparsity

Definition (Recall: sparsity)
A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u, 3s^i)}) \leq q \cdot s^i$.

Observation
Consider $u \in N_i$ for some $i$, and a ball $B := B(u, s^i + h)$ where $h$ is sampled from $[0, s^i]$ uniformly at random. Define $X(h)$ the number of short edges cut by $B$. 

Short Edges: Using Sparsity

Definition (Recall: sparsity)
A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u,3s^i)}) \leq q \cdot s^i$.

Observation
Consider $u \in N_i$ for some $i$, and a ball $B := B(u, s^i + h)$ where $h$ is sampled from $[0, s^i]$ uniformly at random.
Define $X(h)$ the number of short edges cut by $B$. Then

$$\int_0^{s^i} X(h) dh \leq w(F|_{B(u,3s^i)}) \leq q \cdot s^i.$$
Short Edges: Using Sparsity

Definition (Recall: sparsity)
A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u,3s^i)}) \leq q \cdot s^i$.

Observation
Consider $u \in N_i$ for some $i$, and a ball $B := B(u, s^i + h)$ where $h$ is sampled from $[0, s^i]$ uniformly at random.
Define $X(h)$ the number of short edges cut by $B$. Then

$$\int_0^{s^i} X(h) dh \leq w(F|_{B(u,3s^i)}) \leq q \cdot s^i.$$  

So $\Pr[X(h) \leq \frac{q}{2}] \geq \frac{1}{2}$. 

Short Edges: Using Sparsity

Definition (Recall: sparsity)
A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u,3s^i)}) \leq q \cdot s^i$.

Observation
Consider $u \in N_i$ for some $i$, and a ball $B := B(u, s^i + h)$ where $h$ is sampled from $[0, s^i]$ uniformly at random. Define $X(h)$ the number of short edges cut by $B$. Then

$$\int_0^{s^i} X(h)dh \leq w(F|_{B(u,3s^i)}) \leq q \cdot s^i.$$ 

So $\Pr[X(h) \leq \frac{q}{2}] \geq \frac{1}{2}$.

Good Radius
Sparsity implies that for each $u \in N_i$, $B(u, s^i + h_u^{(i)})$ cuts at most $O(q)$ short edges with constant probability.
Short Edges: Using Sparsity

Definition (Recall: sparsity)
A graph $F$ is $q$-sparse, if for all $i$ and $u \in N_i$, $w(F|_{B(u,3s^i)}) \leq q \cdot s^i$.

Observation
Consider $u \in N_i$ for some $i$, and a ball $B := B(u, s^i + h)$ where $h$ is sampled from $[0, s^i]$ uniformly at random.
Define $X(h)$ the number of short edges cut by $B$. Then

$$\int_{0}^{s^i} X(h) dh \leq w(F|_{B(u,3s^i)}) \leq q \cdot s^i.$$ 

So $\Pr[X(h) \leq \frac{q}{2}] \geq \frac{1}{2}$.

Good Radius
Sparsity implies that for each $u \in N_i$, $B(u, s^i + h_u^{(i)})$ cuts at most $O(q)$ short edges with constant probability.
In the following, we condition on such event for all $i$ and $u \in N_i$. 
Limited Cut from Each Height

Suppose $C$ is a cluster centered at $u \in N_i$.

Claim

For each height $j$ ($j \geq i$), there are at most $2^{O(k)}$ height-j clusters cutting a short edge that crosses $C$. 
Limited Cut from Each Height

Suppose $C$ is a cluster centered at $u \in N_i$.

Claim
For each height $j$ ($j \geq i$), there are at most $2^{O(k)}$ height-$j$ clusters cutting a short edge that crosses $C$.

Proof.
Since the height-$j$ clusters are cutting the short edge, their centers are of distance at most $O(s^j + s^i) = O(s^j)$ from $u$. The claim follows from the packing property. \qed
Limited Cut from Each Height

Suppose $C$ is a cluster centered at $u \in N_i$.

**Claim**

For each height $j$ ($j \geq i$), there are at most $2^{O(k)}$ height-$j$ clusters cutting a short edge that crosses $C$.

Therefore, clusters from each height $j$ ($j \geq i$) can contribute $O(q) \cdot 2^{O(k)}$ (short) cuts of $C$, and we denote this number as $Z$. 
Limited Cut from Each Height

Suppose $C$ is a cluster centered at $u \in N_i$.

**Claim**

*For each height $j$ ($j \geq i$), there are at most $2^{O(k)}$ height-$j$ clusters cutting a short edge that crosses $C$.*

Therefore, clusters from each height $j$ ($j \geq i$) can contribute $O(q) \cdot 2^{O(k)}$ (short) cuts of $C$, and we denote this number as $Z$.

**How about $j < i$?**

$C$ is determined by random radius for $u \in N_j$ and $j \geq i$ only.
Short Edges: Better Charging Argument

Set $r$ to be $2 \log_s L \cdot Z + (O\left(\frac{sk}{\epsilon}\right))^k$. 

Limited Cut from Lower Clusters

If $r' > r$, then at most $\log_s L \cdot Z \leq r'^2$ edges are cut by clusters no higher than $(i + \log_s L)$.

New Charging Scheme

So at least $r' - r^2 > r'^2$ edges are cut by clusters higher than $(i + \log_s L)$. Charge the cost to those edges, and each takes $O\left(s^i\right) \cdot r'^{1 - \frac{1}{k}} \leq O\left(s^i\right) \cdot r$. 


Short Edges: Better Charging Argument

Set $r$ to be $2 \log_s L \cdot Z + (O\left(\frac{sk}{\epsilon}\right))^k$.

**Limited Cut from Lower Clusters**

If $r' > r$, then at most $\log_s L \cdot Z \leq \frac{r}{2}$ edges are cut by clusters no higher than $(i + \log_s L)$.
Short Edges: Better Charging Argument

Set $r$ to be $2 \log_s L \cdot Z + (O(\frac{sk}{\epsilon}))^k$.

**Limited Cut from Lower Clusters**

If $r' > r$, then at most $\log_s L \cdot Z \leq \frac{r}{2}$ edges are cut by clusters no higher than $(i + \log_s L)$.

**New Charging Scheme**

So at least $r' - \frac{r}{2} > \frac{r'}{2}$ edges are cut by clusters higher than $(i + \log_s L)$. Charge the cost to those edges, and each takes

$$O(s^i) \cdot \frac{r'^{1 - \frac{1}{k}}}{\frac{r'}{2}} \leq O(s^i) \cdot r^{-\frac{1}{k}}.$$
Short Edges: Better Charging Argument

Accounting: Single Height
At height $i$ clusters, each edge $(u, v)$ takes $O(s^i) \cdot r^{-\frac{1}{k}}$, with probability

$$O(k \frac{d(u, v)}{s^i + \log s L}) = O(k \frac{d(u, v)}{L \cdot s^i}).$$
Short Edges: Better Charging Argument

Accounting: Single Height
At height $i$ clusters, each edge $(u, v)$ takes $O(s^i) \cdot r^{-\frac{1}{k}}$, with probability

$$O(k \frac{d(u, v)}{s^i + \log s \cdot L}) = O(k \frac{d(u, v)}{L \cdot s^i}).$$

Accounting: Over All Heights
Over all heights, the expected cost is at most

$$O(k) \cdot \sum_{i=1}^{L} \frac{d(u, v)}{L \cdot s^i} \cdot s^i \cdot r^{-\frac{1}{k}} = O(k) \cdot r^{-\frac{1}{k}} \cdot d(u, v)$$

$$= O(\epsilon) \cdot d(u, v),$$

recalling $r := 2 \log s \cdot L \cdot Z + (O(\frac{sk}{\epsilon}))^k > (O(\frac{sk}{\epsilon}))^k$. 

New Challenge: The Steiner Forest Problem

Definition (Steiner Forest Problem (SFP))

Suppose \( M(X, d) \) is a metric space.

- **Given:** a collection of \( n \) terminal pairs \( T := \{(s_i, t_i)\}_{i=1}^n \) \((s_i, t_i \in X)\).
- **Goal:** a minimum weight graph (induced by \( M \)) that connects each pair \((s_i, t_i) \in T\).
Definition (Steiner Forest Problem (SFP))

Suppose $M(X, d)$ is a metric space.

- Given: a collection of $n$ terminal pairs $T := \{(s_i, t_i)\}_{i=1}^n$ ($s_i, t_i \in X$).
- Goal: a minimum weight graph (induced by $M$) that connects each pair $(s_i, t_i) \in T$. 
New Challenge: The Steiner Forest Problem

Definition (Steiner Forest Problem (SFP))

Suppose \( M(X, d) \) is a metric space.

- Given: a collection of \( n \) terminal pairs \( T := \{(s_i, t_i)\}_{i=1}^{n} \) \((s_i, t_i \in X)\).
- Goal: a minimum weight graph (induced by \( M \)) that connects each pair \((s_i, t_i) \in T\).
New Challenge: The Steiner Forest Problem

Definition (Steiner Forest Problem (SFP))
Suppose $M(X, d)$ is a metric space.
- Given: a collection of $n$ terminal pairs $T := \{(s_i, t_i)\}_{i=1}^n$ ($s_i, t_i \in X$).
- Goal: a minimum weight graph (induced by $M$) that connects each pair $(s_i, t_i) \in T$.

Non-terminal points are called *Steiner points*.
Special case: Steiner tree.
Difficulty of DP: Encoding Connectivity

The \((m, r)\)-light solution idea still works, but is not sufficient.
Difficulty of DP: Encoding Connectivity

The \((m, r)\)-light solution idea still works, but is not sufficient.

Difficulty

Clusters may separate terminal pairs.

Figure: combining two subproblems: is this feasible?

DP needs to track which portals each terminal connects to.
Difficulty of DP: Encoding Connectivity

The \((m, r)\)-light solution idea still works, but is not sufficient.

Difficulty

Clusters may separate terminal pairs.

DP needs to track which portals each terminal connects to. Naive way is exponential time: \(2^r \Omega(n)\).
Cells and Cell Property

Idea: group several terminals together (each group is called a cell), and make terminals in a cell have similar connectivity [BKM08].
**Cells and Cell Property**

**Idea**: group several terminals together (each group is called a cell), and make terminals in a cell have *similar connectivity* [BKM08]. A possible way of grouping:

**Uniform Cells [BKM08]**

For each cluster $C$, let its sub-clusters of scale $\gamma s^i$ ($0 < \gamma < 1$) be the cells, where $C$ is of height-$i$. 
Cell Property

A solution $F$ satisfies the cell property if for each cell, there is at most one component in $F$ that both crosses $C$ and touches the cell.

(a) Cell Property Violated

(b) Cell Property Satisfied
Incorporating Cell Property in DP

Let $\text{Cel}(C)$ be the cell set of $C$. 
Incorporating Cell Property in DP

Let $\text{Cel}(C)$ be the cell set of $C$.

**Structural Property**

There exists $(m, r)$-light solution $F$ that satisfies the cell property with respect to Cel, such that $w(F) \leq (1 + \epsilon) \cdot \text{OPT}$.
Incorporating Cell Property in DP

Let \( \text{Cel}(C) \) be the cell set of \( C \).

**Structural Property**

There exists \((m, r)\)-light solution \( F \) that satisfies the cell property with respect to \( \text{Cel} \), such that \( w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Cell Property: New Attributes for Subproblems**
Incorporating Cell Property in DP

Let \( \text{Cel}(C) \) be the cell set of \( C \).

**Structural Property**

There exists \((m, r)\)-light solution \( F \) that satisfies the cell property with respect to \( \text{Cel} \), such that \( w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Cell Property: New Attributes for Subproblems**

- Cell set \( \text{Cel} \).
Incorporating Cell Property in DP

Let Cel(C) be the cell set of C.

Structural Property
There exits \((m, r)\)-light solution \(F\) that satisfies the cell property with respect to Cel, such that \(w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

Cell Property: New Attributes for Subproblems

- Cell set Cel.
- A mapping \(g\) from Cel to \(2^R\), denoting the subset of \(R\) that a cell connects to.
Incorporating Cell Property in DP

Let \( \text{Cel}(C) \) be the cell set of \( C \).

**Structural Property**
There exits \((m, r)\)-light solution \( F \) that satisfies the cell property with respect to \( \text{Cel} \), such that \( w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Cell Property: New Attributes for Subproblems**

- Cell set \( \text{Cel} \).
- A mapping \( g \) from \( \text{Cel} \) to \( 2^R \), denoting the subset of \( R \) that a cell connects to.

The mapping \( g \) has \( 2^{|R| \cdot |\text{Cel}(C)|} \leq 2^{r \cdot |\text{Cel}(C)|} \) possibilities.
Incorporating Cell Property in DP

Let \( \text{Cel}(C) \) be the cell set of \( C \).

**Structural Property**

There exists \((m, r)\)-light solution \( F \) that satisfies the cell property with respect to \( \text{Cel} \), such that \( w(F) \leq (1 + \epsilon) \cdot \text{OPT} \).

**Cell Property: New Attributes for Subproblems**

- Cell set \( \text{Cel} \).
- A mapping \( g \) from \( \text{Cel} \) to \( 2^R \), denoting the subset of \( R \) that a cell connects to.

The mapping \( g \) has \( 2^{|R| \cdot |\text{Cel}(C)|} \leq 2^{r \cdot |\text{Cel}(C)|} \) possibilities.

**Can We Use Uniform Cells?**
Incorporating Cell Property in DP

Let $\text{Cel}(C)$ be the cell set of $C$.

**Structural Property**

There exists $(m, r)$-light solution $F$ that satisfies the cell property with respect to $\text{Cel}$, such that $w(F) \leq (1 + \epsilon) \cdot \text{OPT}$.

**Cell Property: New Attributes for Subproblems**

- Cell set $\text{Cel}$.
- A mapping $g$ from $\text{Cel}$ to $2^R$, denoting the subset of $R$ that a cell connects to.

The mapping $g$ has $2^{|R| \cdot |\text{Cel}(C)|} \leq 2^r \cdot |\text{Cel}(C)|$ possibilities.

**Can We Use Uniform Cells?**

No. For doubling metrics, we can achieve $\gamma \approx \frac{\epsilon}{\log n}$ only, which implies $|\text{Cel}(C)| \approx \left(\frac{\log n}{\epsilon}\right)^k$. $2^r |\text{Cel}(C)|$ is quasi-poly.
Our Technique: Adaptive Cells

Intuition

(a) Uniform Cells

(b) Adaptive Cells

\((m, r)\)-light implies there are at most \(r\) components crossing \(C\) for any cluster \(C\).

We only need cells for which crossing components touch.

We use cells of adaptive scales.
Our Technique: Adaptive Cells

Intuition

(a) Uniform Cells

(b) Adaptive Cells

- $(m, r)$-light implies there are at most $r$ components crossing $C$ for any cluster $C$. 
Our Technique: Adaptive Cells

Intuition

(a) Uniform Cells
(b) Adaptive Cells

- \((m, r)\)-light implies there are at most \(r\) components crossing \(C\) for any cluster \(C\).
- We only need cells for which crossing components touch.
Our Technique: Adaptive Cells

Intuition

- (m, r)-light implies there are at most r components crossing $C$ for any cluster $C$.
- We only need cells for which crossing components touch.
- We use cells of adaptive scales.
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

Adaptive Cells

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\begin{cases}
  s_i \text{ if } w(A) \geq s_i \\
  s_i \log n \leq w(A) < s_i \\
  s_i \log n \text{ if } w(A) < s_i
\end{cases}
\]

Cells are of scale \([s_i \log n, s_i]\) (\(O(\log \log n)\) scales).

How many cells are created for \(A\)?
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).
Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

**Adaptive Cells**

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\left\{
\begin{array}{ll}
\text{if } w(A) \geq s_i \\
\text{if } s_i \log n \leq w(A) < s_i \\
\text{if } w(A) < s_i
\end{array}
\right.
\]

Cells are of scale \([s_i \log n, s_i]\) (\(O(\log \log n)\) scales). How many cells are created for \(A\)?
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

Adaptive Cells

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\begin{cases} 
    s^i & \text{if } w(A) \geq s^i \\
\end{cases}
\]

Cells are of scale \([s^i, s^i]\) (\(O(\log \log n)\) scales).

How many cells are created for \(A\)?
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

Adaptive Cells

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\begin{cases}
  s^i & \text{if } w(A) \geq s^i \\
  \frac{s^i}{\log n} & \text{if } \frac{s^i}{\log n} \leq w(A) < s^i
\end{cases}
\]
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

Adaptive Cells

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\begin{cases} 
  s^i & \text{if } w(A) \geq s^i \\
  w(A) & \text{if } \frac{s^i}{\log n} \leq w(A) < s^i \\
  \frac{s^i}{\log n} & \text{if } w(A) < \frac{s^i}{\log n}
\end{cases}
\]
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

Adaptive Cells

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\begin{align*}
&\frac{s^i}{\log n} & \text{if } w(A) < \frac{s^i}{\log n} \\
&w(A) & \text{if } \frac{s^i}{\log n} \leq w(A) < s^i \\
&s^i & \text{if } w(A) \geq s^i
\end{align*}
\]

Cells are of scale \([\frac{s^i}{\log n}, s^i]\) \((O(\log \log n)\) scales).
An Example Implementation of Adaptive Cells

Fix a \((m, r)\)-light solution \(F\). Consider a cluster \(C\) of height \(i\).

Adaptive Cells

For each crossing component \(A\), define the adaptive cells of \(A\) to be the sub-clusters of \(C\) that intersect \(A\) with scale

\[
\begin{align*}
    s^i & \quad \text{if } w(A) \geq s^i \\
    w(A) & \quad \text{if } \frac{s^i}{\log n} \leq w(A) < s^i \\
    \frac{s^i}{\log n} & \quad \text{if } w(A) < \frac{s^i}{\log n}
\end{align*}
\]

Cells are of scale \([\frac{s^i}{\log n}, s^i]\) (\(O(\log \log n)\) scales).

How many cells are created for \(A\)?
Typical Case: \( \frac{s^i}{\log n} \leq w(A) < s^i \)

Recall that the scale of cells in this case is \( \approx w(A) \).
Typical Case: \( \frac{s^i}{\log n} \leq w(A) < s^i \)

Recall that the scale of cells in this case is \( \approx w(A) \).

- Bounded Diameter. The diameter of all adaptive cells of \( A \) is \( O(w(A)) \).
Typical Case: \( \frac{s^i}{\log n} \leq w(A) < s^i \)

Recall that the scale of cells in this case is \( \approx w(A) \).

- **Bounded Diameter.** The diameter of all adaptive cells of \( A \) is \( O(w(A)) \).
- **Packing.** Moreover, the centers of these cells form a \( \Omega(w(A)) \)-packing.
Typical Case: \( \frac{s^i}{\log n} \leq w(A) < s^i \)

Recall that the scale of cells in this case is \( \approx w(A) \).

- **Bounded Diameter.** The diameter of all adaptive cells of \( A \) is \( O(w(A)) \).
- **Packing.** Moreover, the centers of these cells form a \( \Omega(w(A)) \)-packing.

By packing property, the number of adaptive cells for \( A \) is at most \( O(1)^k \). We can show this similarly for the other two cases.
Typical Case: $\frac{s^i}{\log n} \leq w(A) < s^i$

Recall that the scale of cells in this case is $\approx w(A)$.

- **Bounded Diameter.** The diameter of all adaptive cells of $A$ is $O(w(A))$.
- **Packing.** Moreover, the centers of these cells form a $\Omega(w(A))$-packing.

By packing property, the number of adaptive cells for $A$ is at most $O(1)^k$. We can show this similarly for the other two cases.

$(m, r)$-light implies $|\text{Cel}(C)| \leq r \cdot O(1)^k$. 
Counting New Attributes

Recall we use sparsity and

\[ r = \log^c(n) \]

for small constant \( c \in (0, 1) \).
Counting New Attributes

Recall we use sparsity and

\[ r = \log^c(n) \]

for small constant \( c \in (0, 1) \).

Counting Cel: Without the Knowledge of \( F \)
Counting New Attributes

Recall we use sparsity and

\[ r = \log^c(n) \]

for small constant \( c \in (0, 1) \).

Counting Cel: Without the Knowledge of \( F \)

Recall that cells are of scale \([\frac{s^i}{\log n}, s^i]\) for a height-\( i \) cluster.

The number of candidate cells: \( \approx (\log n)^k \).
Counting New Attributes

Recall we use sparsity and

\[ r = \log^c(n) \]

for small constant \( c \in (0, 1) \).

Counting Cel: Without the Knowledge of \( F \)

Recall that cells are of scale \([\frac{s^i}{\log n}, s^i]\) for a height-\( i \) cluster.

The number of candidate cells: \( \approx (\log n)^k \).

So the number of possible Cel is

\[
\left( \frac{(\log n)^k}{O(1)^k \cdot r} \right) \leq O(\log n)^O(r) = 2^{\log^c+o(1)(n)}.
\]
Counting New Attributes

Recall we use sparsity and

\[ r = \log^c(n) \]

for small constant \( c \in (0, 1) \).

Counting Cel: Without the Knowledge of \( F \)

Recall that cells are of scale \( \left[ \frac{s^i}{\log n}, s^i \right] \) for a height-\( i \) cluster.

The number of candidate cells: \( \approx (\log n)^k \).

So the number of possible Cel is

\[
\binom{(\log n)^k}{O(1)^k \cdot r} \leq O(\log n)^O(r) = 2^{\log^c \alpha(1)(n)}.
\]

Counting the Mapping \( g \)

The possibilities of mapping \( g \) from \( \text{Cel}(C) \) to \( 2^R \) is (recalling that \( |\text{Cel}(C)| \leq r \cdot O(1)^k \))

\[
2^{r \cdot |\text{Cel}(C)|} \leq 2^{O(\log^{2c}(n))}.
\]
Open Question: Prize Collecting TSP

Is there a PTAS for the prize-collecting TSP in doubling metrics?
Open Question: Prize Collecting TSP

Is there a PTAS for the prize-collecting TSP in doubling metrics?

Known in Euclidean Space

A PTAS for the prize-collection TSP in Euclidean space is given by Arora [Aro96].
Open Question: Prize Collecting TSP

Is there a PTAS for the prize-collecting TSP in doubling metrics?

Known in Euclidean Space
A PTAS for the prize-collection TSP in Euclidean space is given by Arora [Aro96].

PTAS?
The current framework is not readily to be applied, especially the reduction to sparse instances may be difficult.
Open Question: Prize Collecting TSP

Is there a PTAS for the prize-collecting TSP in doubling metrics?

**Known in Euclidean Space**
A PTAS for the prize-collection TSP in Euclidean space is given by Arora [Aro96].

**PTAS?**
The current framework is not readily to be applied, especially the reduction to sparse instances may be difficult.

**APX-Hardness?**
Is this problem a separation between Euclidean spaces and doubling metrics?
Thank you!


Sanjeev Arora, Prabhakar Raghavan, and Satish Rao.
Approximation schemes for euclidean k-medians and related problems.

Patrice Assouad.
Plongements lipschitziens dans $\mathbb{R}^n$.

Yair Bartal, Lee-Ad Gottlieb, and Robert Krauthgamer.
The traveling salesman problem: low-dimensionality implies a polynomial time approximation scheme.

MohammadHossein Bateni and MohammadTaghi Hajiaghayi.
Euclidean prize-collecting steiner forest.

Glencora Borradaile, Philip N. Klein, and Claire Mathieu.
A polynomial-time approximation scheme for euclidean steiner forest.
Miroslav Chlebík and Janka Chlebíková.
The steiner tree problem on graphs: Inapproximability results.

T.-H. Hubert Chan and Khaled M. Elbassioni.
A QPTAS for TSP with fat weakly disjoint neighborhoods in
doubling metrics.

A ptas for the steiner forest porblem in doubling metrics.

T.-H. Hubert Chan and Shaofeng H.-C. Jiang.
Reducing curse of dimensionality: Improved PTAS for TSP
(with neighborhoods) in doubling metrics.

Michel X. Goemans and David P. Williamson.
A general approximation technique for constrained forest problems.


**Kunal Talwar.**

Bypassing the embedding: algorithms for low dimensional metrics.