Geometric Approximations:
- Useful when exact computation is too costly
- Geometric inputs are "measurements" and often are uncertain. So approximate solutions are fine.

Examples:
Euclidean MST of pt set \( P = \{ p_1, \ldots, p_n \} \subseteq \mathbb{R}^d \)

Exact: \( O(n \log n) \) in \( \mathbb{R} \)
\( O(n^{2-\frac{4}{d}}) \) in \( \mathbb{R}^d \) [Nearly quadratic]

Approx: Given \( \epsilon > 0 \), compute a spanning tree of weight \( \leq (1+\epsilon) \cdot EMST(P) \)
Convex Hull of a set \( P = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d \):

- **Exact:** \( O(n \log n) \) in \( \mathbb{R} \)
- \( O(n^{2d/d}) \) in \( \mathbb{R}^d \)

Approx: Compute a subset \( P' \subseteq P \) s.t. \( \text{conv}(P) \) and \( \text{conv}(P') \) are very similar

Well-Separated Pair Decomposition:

Given set \( P = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d \), the Euclidean graph is complete graph on \( P \), where \( \omega(p_i, p_j) = \|p_i - p_j\| \)

- Has \( \binom{n}{2} = O(n^2) \) edges
- Can we encode this using a structure of size \( O(n) \)?
**Intuition:** If two point clusters $A, B \subseteq P$ are well separated, we can represent many edges of $A \times B$ using a single edge connecting a representative $a \in A$ and $b \in B$.

If we do this for all well-separated clusters, how many edges do we need?

**Def:** Given $P = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d$ and scalar $s > 0$,

- Two sets $A, B \subseteq P$ are *s-well-separated* if $A + B$ can be enclosed in balls of some radius $r$, such that these balls are separated by distance $\geq s \cdot r$.
**Obs:**
- If $A + B$ are $s$-well separated, they are $s'$-well separated for any $0 < s' < s$.

- Two singleton sets $A = \{a\}$, $B = \{b\}$ are $s$-well separated for any $s > 0$. ($a \neq b$)

**Def.** Given sets $A, B$, define

$$A \otimes B = \{\{a, b\} \mid a \in A, b \in B, a \neq b\}$$

**Obs:** $P \otimes P = \text{set of all } (n \choose 2) \text{ pairs of } P$. 
Definition: Given \( P + s > 0 \), an \textit{s-well separated pair decomposition} of \( P \) (s-WSPD) is a collection of pairs
\[
\{ \{ A_1, B_1 \}, \ldots, \{ A_m, B_m \} \}
\]
such that:

1. \( A_i, B_i \subseteq P \) for \( 1 \leq i \leq m \)
2. \( A_i \cap B_i = \emptyset \) (disjoint)
3. \( \bigcup_{i=1}^{m} A_i \cap B_i = P \times P \) (cover)
4. \( A_i + B_i \) are \textit{s-well separated} for \( 1 \leq i \leq m \)

28 pairs

\[ \{\{a,b\}, \{e,f,g,3\}\} \]
Obs: For any $s > 0$ there is always a trivial $s$-WSPPD consisting of $\binom{n}{2}$ singleton pairs.

Can we do better? $d$ is constant; hidden $d$

Yes! $\Rightarrow O(s^d \cdot n)$ pairs

If $s, d$ constants: $O(n)$!
Can compute in time:
$O(n \log n + s^d n)$

Quadtrees:
A tree storing $\mathcal{P}$ based on recursive subdiv.
into hypercubes.

- Let $C_0$ be a bounding hypercube for $\mathcal{P}$
- While a cell of subdivision has 2 or more pts of $\mathcal{P}$, split it into $2^d$ hypercubes of half side length
Note: A quadtree may have more than $O(n)$ nodes, but we can reduce storage to $O(n)$ by path compression. (see latex notes)

Thm: Given a set of pts $P = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d$, we can construct a (compressed) quadtree of space $O(n)$ in $O(n \log n)$ time.

Additional information (provided by construction)

Given node $u$ in tree:

- $\text{level}(u) =$ level of $u$ in tree
- $P(u) =$ set of pts in $u$'s subtree
- $\text{rep}(u) =$ an arbitrary element of $P(u)$

We will represent each WSP as a pair of nodes $\{u, v\}$. Actual pair is $\{P(u), P(v)\}$.
Constructing the WSPD:

Given $P + s > 0$:
- Build **quadtree for $P$** → Let $u_0$ - root
- Invoke: $ws$-pairs$(u_0, u_0, s)$

```
ws-pairs(Node u, Node v, Scalar s) {
    if (u + v are both leaves + u == v) return φ
    if (rep(u) or rep(v) is empty) return φ
    else if (u + v are s-well sep)
        return {u, v} // wsp = {P(u), P(v)}
    else // not ws.
        if (level(u) > level(v))
            swap u ↔ v // u is not deeper than v
        let $u_1, ..., u_k$ be u’s children
        return $\bigcup_{i=1}^k ws$-pairs$(u_i, v, s)$
}
```

Cases: $u + v$ are well sep

$u + v$ not well-sep
Analysis: We'll show \(O(s^d \cdot n)\) pairs generated.

- Assume: Quadtree is not compressed (simpler)
  \[s \geq 1\] (else just use \(s' = \max(1, s)\))

1. Terminal / Non-Terminal:
   - To count no. of WSP's, we'll count no. of calls to ws-pairs.
   - A call is:
     - terminal: makes no recursive calls
     - non-terminal: otherwise
   - It suffices to count just no. of non-terminal calls (each generates at most \(2^d = O(1)\) term. calls)

2. Charging: We'll count no. of non-term calls by charging each to node of tree.
   - Preview: Each node receives \(O(s^d)\) charges.
   - \(O(n)\) nodes in tree
   \[\Rightarrow O(s^d \cdot n)\] total charges
Let \( \text{ws-pairs} (u, v, s) \) be non-term call

\[ \Rightarrow u, v \text{ not well sep.} \]

\[ \Rightarrow \text{Assume (w.l.o.g.) } \text{lev}(u) \leq \text{lev}(v) \]

\[ \Rightarrow \text{we will charge } v \]

(smaller node is charged)

- Let \( x \) be side length of \( v \)'s cell
- We always split larger cell first

\[ \Rightarrow u \text{'s side length } \leq 2x \]

- Let \( r_v = \text{radius of ball enclosing } v \text{'s cell} \)
  
  \[ r_u = 2r_v \]

\[ \Rightarrow r_u \leq 2r_v \]

and \( r_v = x \sqrt{2}/2 \)

- Let \( c_u, c_v \) be centers of \( u \) and \( v \) cells
This cell is non-term
⇒ u, v not well separated
⇒ Distance between balls is
  \( s \cdot \max(r_u, r_v) \leq s \cdot r_u \leq s(2 \cdot r_v) \)
  \( = s \cdot x \cdot \sqrt{d} \)
⇒ Distance between centers
  \[ \|c_u - c_v\| \leq r_v + r_u + s \cdot x \sqrt{d} \]
  \( \leq x \sqrt{d}/2 + x \sqrt{d} + s \cdot x \sqrt{d} \)
  \( = (\frac{1}{2} + 1 + s) x \sqrt{d} \)
  \( < 3 s x \sqrt{d} \) (since \( s \geq 1 \))

Def: \( R_v = 3 s x \sqrt{d} \)

Summary: A node \( v \) of side length \( x \)
is charged by nodes \( u \) of side length \( x \) or \( 2x \) whose cell centers lie within a ball of radius \( R_v = 3 s x \sqrt{d} \) of \( c_v \).

How many such nodes can there be?
Packing Lemma: Given a ball $b$ of radius $r$ in $\mathbb{R}^d$ and any collection $X$ of disjoint quadtree cells of side length $\geq x$ that overlap $b$, then

$$|X| \leq (1 + \left\lceil \frac{2r}{x} \right\rceil)^d \leq O(\max(2, \frac{r}{x})^d)$$

Proof: To maximize no. of cells, assume they are as small as possible $\Rightarrow x$

These cells form a grid of side length $x$ that overlaps $b$

\[ \text{No. of grid squares of side length } x \text{ overlapping an interval of length } 2r \text{ is} \]

\[ \leq 1 + \left\lceil \frac{2r}{x} \right\rceil \]

$\Rightarrow \text{Total:} (1 + \left\lceil \frac{2r}{x} \right\rceil)^d$
Returning to WSPD analysis:

- No. of charges to $v$ ≤

$$\text{No. of nodes of side length } \geq x \text{ overlapping a ball of radius } R_v = 3s \times 1/2$$

- By Packing Lemma, no. of nodes

$$\leq (1 + \left\lceil \frac{2R_v}{x} \right\rceil)^d$$

$$\leq (1 + \left\lceil \frac{6s \times 1/2}{x} \right\rceil)^d$$

$$\leq (2 + 6s \sqrt{d})^d$$

$$\leq O(s^d) \text{ since } s > 1$$

So, each node charged $O(s^d)$ times

- $O(n)$ nodes in quadtree

- $O(n \cdot s^d)$ non-term calls to ws-pairs

- $O(n \cdot s^d)$ pairs generated

whew!!
Theorem: Given a point set $P = \{p_1, \ldots, p_n\}$ in $\mathbb{R}^d$ ($d$ is constant) and $s \geq 1$, in $O(n \log n + s^d n)$ time, can build an $s$-WSPD for $P$ of size $O(s^d n)$.