# Applications of the WSPD

Anne Driemel and Herman Haverkort updated: January 10, 2025

In this lecture we will discuss several algorithmic applications of the well-separated pair decomposition (WSPD).

## 1 Computing a Spanner

It is sometimes useful to capture the distances between the $\binom{n}{2}$ pairs in a set of $n$ points by the shortest-path distances in a sparse graph that contains only few edges. For this, we define the notion of spanners.

**Definition 19.1** ($t$-Spanner)**.** *A spanner for a set of points $P$ is a connected graph $G$ whose vertices are the points of $P$. By $d_G(p, q)$ we denote the minimum weight of any path from p to q in $G$ (the shortest path). A spanner $G$ is a $t$-spanner for $P$ if the following holds for any points $p, q \in P$: $d_G(p, q) \leq t \cdot \|p - q\|$.*

To obtain a $t$-spanner, we build a compressed quadtree, and then construct a WSPD with separation ratio $s = 4(t + 1)/(t - 1)$. The spanner now consists of one edge for each pair of the WSPD: the edge for the pair $\{A_i, B_i\}$ connects an arbitrary point $a_i \in A_i$ to an arbitrary point $b_i \in B_i$. We can construct such a spanner in time linear in the size of the WSPD, if we assume that each node in the quadtree stores a direct pointer to one of its descendant leaves that stores a point of $P$. This can be ensured with a simple post-order traversal of the compressed quadtree in linear time. Perhaps surprisingly, the resulting graph is connected. We will show correctness in the following theorem.

**Theorem 19.2.** *Given a compressed quadtree storing a set $P$ of $n$ points, and a value $\varepsilon > 0$, we can compute a $(1 + \varepsilon)$-spanner of the points in $O(n/\varepsilon^d)$ time. The resulting spanner has $O(n/\varepsilon^d)$ edges.*

*Proof.* Let $G$ be the graph computed by the above algorithm, for a value of $t$ that we will specify later. We will prove for any $p, q \in P$, that we have $d_G(p, q) \leq t \cdot \|p - q\|$, by induction on the rank of $\|p - q\|$ in the list of pairwise distances in $P$, sorted in increasing order. (This also shows that the graph is connected). Clearly, the claim holds for rank zero, that is, $\|p - q\| = 0$. Now, consider any pair of distinct points $p, q$. There is at least one pair $\{A_i, B_i\}$ that covers $\{p, q\}$; without loss of generality we assume $p \in A_i$ and $q \in B_i$. Let $\{a_i, b_i\}$ be the corresponding edge in the spanner. Because $\{A_i, B_i\}$ is a well-separated pair and $s > 2$, we have $\|p - a_i\| \leq (2/s)\|p - q\| < \|p - q\|$, so $\|p - a_i\|$ has lower rank than $\|p - q\|$ and therefore, by induction, we have

$$d_G(p, a_i) \leq t \cdot \|p - a_i\|$$

By an analogous argument, we have

$$d_G(b_i, q) \leq t \cdot \|b_i - q\|$$

Furthermore, because $\{A_i, B_i\}$ is a well-separated pair, $\|a_i - b_i\| \leq (1 + 4/s)\|p - q\|$. Therefore we have:

$$\begin{aligned} d_G(p,q) &\leq d_G(p,a_i) + d_G(a_i,b_i) + d_G(b_i,q) \quad \text{(by triangle inequality)} \\ &\leq t \cdot \|p-a_i\| + \|a_i - b_i\| + t \cdot \|b_i - q\| \\ &\leq \left(\frac{2t}{s}\right)\|p-q\| + \left(1+\frac{4}{s}\right)\|p-q\| + \left(\frac{2t}{s}\right)\|p-q\| \\ &= \left(1+\frac{4(1+t)}{s}\right)\|p-q\| \\ &= t\cdot\|p-q\| \quad \text{(by our choice of } s\text{)} \end{aligned}$$

It follows that $G$ is a $t$-spanner with $O(s^d n)$ edges. If we choose $t = 1+\varepsilon$, then $s = 4(t+1)/(t-1) = 4(2+\varepsilon)/\varepsilon = 8/\varepsilon + 4$, the spanner has $O(n/\varepsilon^d)$ edges. □

## 2 Minimum spanning tree

**Definition 19.3** (Spanning trees)**.** *A spanning tree of a set of points $P$ is a tree structure of which the vertices are exactly the points of $P$. The weight of a spanning tree is the total length of its edges. A (geometric) minimum spanning tree of $P$ is a spanning tree of $P$ of lowest possible weight.*

A minimum spanning tree of a set $P$ of $n$ points in $d$-dimensional Euclidean space can be computed by first computing a complete graph that has a vertex for each point in $P$ and where an edge $(p,q)$ has weight $\|p-q\|$. This requires computing all pairwise distances between points in $P$ and then applying an algorithm to compute the minimum spanning tree. Kruskal's MST-algorithm, for example, has running time in $O(m \log m)$, where $m$ is the number of edges. So, this naive approach would result in a running time of $O(n^2 \log n)$.

Can we do faster? For $d = 2$, the MST is a subgraph of the Delaunay triangulation[1], which reduces the number of edges that need to be considered from quadratic to linear. However, when $d > 2$, this does not work anymore, since the complexity of the Voronoi diagram (and therefore also its dual, the Delaunay graph) grows very fast with $d$. So, this does not seem to be a viable option. Instead, we will see, that we can use WSPD's to obtain at least an approximation to the minimum spanning tree within reasonable running time.

**Definition 19.4** (Approximation MST)**.** *A $t$-approximate minimum spanning tree is a spanning tree with weight as most $t$ times the weight of a true minimum spanning tree.*

The key to obtaining an approximate minimum spanning tree fast is that we first construct a $t$-spanner on the points of $P$. Given a $t$-spanner $G$ of $P$ with $O(n)$ edges, we compute an (exact) minimum spanning tree of $G$ and the result will be a $t$-approximate minimum spanning tree of $P$. The difference is, that we now run the MST-algorithm on a graph with fewer edges.

**Theorem 19.5.** *Given a compressed quadtree storing a set $P$ of $n$ points, and a value $\varepsilon > 0$, we can compute a $(1+\varepsilon)$-approximate minimum spanning tree of the points in $O(n/\varepsilon^d \cdot \log(n/\varepsilon))$ time.*

*Proof.* Consider a minimum-spanning tree $T$ of $P$ and a $t$-spanner $G$ of $P$. Let $G'$ be a graph that consists of, for every edge $(p,q)$ of $T$, a shortest path from $p$ to $q$ in $G$. Since $G$ is a

[1] This follows directly by applying Kruskal's algorithm for computing the MST, since the algorithm in each step considers the shortest edge across a cut between two disjoint vertex sets. Such a shortest edge is always an edge between two neighboring Voronoi regions, and therefore it is a Delaunay edge.

$t$-spanner, the weight of each such path is at most $t \cdot \|p - q\|$, and thus, the weight of $G'$ is at most $t$ times the weight of $T$. Therefore, any spanning tree of $G'$ is a $t$-approximate minimum spanning tree of $P$. Moreover, any such spanning tree of $G'$ is also a spanning tree of $G$. The weight of $\mathrm{MST}(G)$ can only be smaller or equal to the weight of this spanning tree, so its weight is at most $t$ times the weight of $T$. □

## 3 Diameter

The diameter of a set of points $P$ is the largest distance between any two points $p, q \in P$. Naturally, we can use the a WSPD to compute a good approximation of the diameter of a set of points.

**Theorem 19.6.** *Given a compressed quadtree storing a set $P$ of $n$ points, and a value $\varepsilon > 0$, we can compute in $O(n/\varepsilon^d)$ time a pair of points $p, q$ from $P$, such that*

$$\|p - q\| \geq (1 - \varepsilon) \max_{p,q \in P} \|p - q\|$$

(Proof → Exercise)

## 4 Closest pair

A WSPD for a set of points $P \in \mathbb{R}^d$ with separation ratio $s > 2$ has the following property:

**Lemma 19.7.** *If $q$ is a point of $P$ that is closest to a point $p \in P$ among all points in $P \setminus \{p\}$, then any WSPD with separation ratio greater 2 contains a pair $\{\{p\}, B\}$ such that $q \in B$.*

*Proof.* Assume for the sake of contradiction that the statement is false. Then, $\{p, q\}$ must be covered by a pair $\{A, B\}$ where $A$ contains $p$ and at least one other point $p'$. Then we have

$$\|p - p'\| \leq (2/s)\|p - q\| < \|p - q\|$$

so $q$ cannot be the point of $P$ that is closest to $p$ among all points in $P \setminus \{p\}$. □

An immediate consequence is: if $\{p, q\}$ is a closest pair in $P$ (that is a pair of points that realizes $\min_{p,q \in P} \|p - q\|$), then the WSPD contains a pair $\{\{p\}, \{q\}\}$. So, given a linear-size WSPD of $P$, we can find such a pair in $O(n)$ time by checking all singleton pairs in the WSPD and return the one with smallest distance. Note that this is an exact answer, not an approximation.

A WSPD can be computed from a compressed quadtree in linear time (if the dimension $d$ is constant), and the compressed quadtree can be computed in $O(n \log n)$ under certain assumptions (for example, if the coordinates have limited precision, or if we use a real RAM extended with logarithm and rounding operators that run in constant time). Alternatively, the WSPD can be computed from a so-called *fair-split* tree, which can be computed in $O(n \log n)$ time—the interested reader may check out the details in the book by Narasimhan and Smid.

## References


- Giri Narasimhan and Michiel Smid: *Geometric spanner networks*, Cambridge University Press, 2007.
- Sariel Har-Peled: *Geometric Approximation Algorithms*, Mathematical Surveys and Monographs, Volume 173, American Mathematical Society, 2011.