# Approximating Spanning Trees with Low Crossing Number

Sariel Har-Peled*

January 25, 2013

**Abstract**

We present a linear programming based algorithm for computing a spanning tree $T$ of a set $P$ of $n$ points in $\mathbb{R}^d$, such that its crossing number is $O(\min(cr_{\mathrm{span}}(P) \log n, n^{1-1/d}))$, where $cr_{\mathrm{span}}(P)$ the minimum crossing number of any spanning tree of $P$. This is the first guaranteed approximation algorithm for this problem. We provide a similar approximation algorithm for the more general settings of building a spanning tree for a set system with bounded VC dimension.

Our approach is an alternative to the reweighting technique previously used in computing such spanning trees.

Some of the results on this writeup are already known. See introduction for details, see Remark $1.1_{\mathrm{p3}}$.

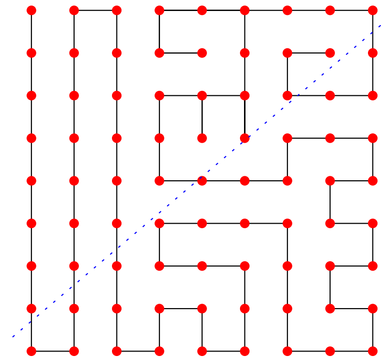
# 1 Introduction

The reweighting technique is a powerful tool in computer science [AHK06]. In Computational Geometry, it was introduced by Chazelle and Welzl [CW89] who used it to compute spanning paths with low crossing number in set systems with bounded VC dimension. Welzl [Wel92] provided a tighter analysis for the case of spanning tree of points in $\mathbb{R}^d$. Matoušek [Mat92] used the reweighting technique to provide a powerful partition theorem that proved to be very useful in building range searching data-structures [AE98]. Also, Clarkson [Cla93] provided an algorithm for polytope approximation that used the reweighting technique. Brönnimann and Goodrich [BG95] realized that Clarkson's algorithm implies a general method for solving hitting set and set cover problems in geometric settings.

Interestingly, Long [Lon01] had observed that set cover problems in geometric settings can be solved by using LP and taking a random sample (guided by the LP solution) that is an $\varepsilon$-net (a similar observation was later made by [ERS05]). In fact, such packing/covering LPs can be solved efficiently via reweighting [PST91]. Thus, one can interpret the reweighting algorithm for solving the geometric set cover problem as directly solving the associated LP.

*Department of Computer Science; University of Illinois; 201 N. Goodwin Avenue; Urbana, IL, 61801, USA; `sariel@uiuc.edu`; `http://www.uiuc.edu/~sariel/`.

The result of Welzl [Wel92], mentioned above, is quite intriguing. It shows that for a set $P$ of $n$ points in the plane (resp., in $\mathbb{R}^d$) one can find a spanning tree of the points, such that any line (resp., hyperplane) crosses at most $O(\sqrt{n})$ (resp., $O(n^{1-1/d})$) edges (i.e., segments) of the spanning tree. To appreciate this result, consider the point set formed by the grid $\sqrt{n} \times \sqrt{n}$. It is trivial in this case to come up with a spanning tree with a crossing number $O(\sqrt{n})$ – any spanning tree of the grid points using only edges of the grid has this property, see figure on the right. Surprisingly, the result of Welzl [Wel92] implies that any point set behaves like a grid point set as far as the crossing number of the optimal spanning tree.



In this work, we establish a connection between computing spanning trees with low crossing number and LPs (i.e., linear programs); that is, we show that spanning trees with low crossing number can be computed using LP rounding.

**Approximate spanning tree with the lowest crossing number.** Given a set system $\mathcal{I} = (P, \mathcal{F})$ of finite VC dimension $\tau$, we show how to compute, in polynomial time[1], a spanning tree of $P$ with crossing number $O(t \log n)$ (assuming $t = \Omega(\log n)$), where $t$ is the minimal crossing number of any spanning tree of $P$. This is done by recursively solving a LP relaxation and rounding it. See Section 3 for details.

Naturally, this algorithm also applies to the Euclidean case. Specially, given a set $P$ of $n$ points in $\mathbb{R}^d$ one can compute, in polynomial time, a spanning tree $T$ such that every hyperplane crosses at most $O(cr_{\mathrm{span}}(P) \log n)$ edges of $T$, where $cr_{\mathrm{span}}(P)$ is the minimum crossing number of any spanning tree of $P$.

Surprisingly, this is the first guaranteed approximation algorithm known for this problem. In particular, achieving such an approximation is mentioned as open in the Open Problems Project (see `http://maven.smith.edu/~orourke/TOPP/P20.html#Problem.20`).

**Spanning trees in $\mathbb{R}^d$ with $O(n^{1-1/d})$ crossings.** We also modify the analysis of our algorithm (but not the algorithm itself) so that it yields worst case bound on the crossing number. Specifically, we get a polynomial time algorithm that, for a given set $P$ of $n$ points in $\mathbb{R}^d$, computes a spanning tree $T$ of $P$ such that any hyperplane in $\mathbb{R}^d$ crosses at most $O(n^{1-1/d})$ edges of $T$. Our proof of the correctness of the algorithm is self contained (except for a relatively easy lemma, see Lemma 4.2), and uses LP duality. We believe the new proof provides a new insight into why such trees exist. In particular, Chazelle and Welzl [CW89] and Welzl [Wel92] proofs of the existence of such spanning trees are simple but somewhat "mysterious" (at least for the author, but other people might not see the mystery).

Here is a sketch of the resulting argument why such trees exist: In the plane, it is sufficient to find spanning forest that span at least $\Omega(n)$ vertices of $P$ (in connected components that are not singletons) and has crossing number $t$. One can find such a (fractional) spanning graph by doing LP relaxation. The dual LP then asks (intuitively) to separate the given $n$

[1] We make the standard assumption that solving a LP of polynomial size takes polynomial time.

points into singletons by a set of lines of minimum cardinality (i.e., for any $\mathsf{p}, \mathsf{q} \in P$, there exists a selected line that crosses the segment $\mathsf{pq}$). It is not hard to show that any such set of lines need to be of size $\Omega(\sqrt{n})$. A somewhat more involved argument (since we are dealing with a fractional solution of an $\mathsf{LP}$ that has some other constraints) implies that the dual $\mathsf{LP}$ is feasible for $t = \sqrt{n}$ and its optimal solution is bounded from below. It follows that the primal $\mathsf{LP}$ is feasible. Now, solving the primal $\mathsf{LP}$ and using a straightforward rounding implies that one can compute the required spanning graph. Applying this recursively by selecting a vertex from each connected component, and overlaying the resulting spanning graphs together results in a connected graph of $P$ with crossing number $O(\sqrt{n})$. See Section 4 for details.

Interestingly, while the above algorithm works for any point set in $\mathbb{R}^d$, one can do slightly better in the planar case, and get a deterministic rounding scheme, see Section 4.1 for details.

**Previous work.** Fekete *et al.* [FLM08] suggested using $\mathsf{LP}$ relaxation to compute a spanning tree with low crossing number. Their $\mathsf{LP}$ is considerably more elaborate than ours (considering all cuts), and their iterated rounding scheme seems to perform quite well in practice (although they are unable to provide a theoretical guarantee on the performance). Furthermore, they prove that computing the spanning tree with minimal crossing number is NP-Hard.

**Remark 1.1.** It turns out one can do better than the results stated in this paper by being more careful and clever about the rounding [CV09]. Furthermore, it turn out that the result was essentially already known [BGRS04]. Nevertheless, some of the results (like the one in Section 4) seems to be new and they might be of interest.

**Organization.** In Section 2 we define some of the quantities we are interested in. In Section 3 we show the $O(t \log n)$ approximation algorithm for spanning trees with low crossing number, for the general case of a set system with low $\mathsf{VC}$ dimension. In Section 4, we specialize this algorithm for the case of points and hyperplanes in $\mathbb{R}^d$. In Section 5, we show two results related to the problem of computing the triangulation of a planar point set with a low crossing number. We discuss our results and some related open problems in Section 6.

# 2 Preliminaries

For a set $\mathcal{F}$ of objects in the plane, we denote by $\mathrm{cr}(\mathcal{F})$ the maximum number of objects of $\mathcal{F}$ that are interested by any line of $P$. This is the ***crossing number*** of $\mathcal{F}$. To this end, we will consider a spanning tree to be a set of its edges (i.e., segments), and a triangulation to be a set of triangles.

Given a point set $P$, we will be interested in the following quantities related to $P$.

**Definition 2.1.** The ***matching crossing number*** $cr_{\mathrm{match}}(P)$ is the minimum crossing number of any prefect matching of $P$.

The ***spanning crossing number*** $cr_{\mathrm{span}}(P)$ is the minimum crossing number of any spanning tree of $P$.

The ***triangulation crossing number*** $cr_{\triangle}(P)$ is the minimum crossing number of any triangulation of $P$.

The ***Steiner triangulation crossing number*** $cr_{\triangle}^{St}(P)$ is the minimum crossing number of any Steiner triangulation of $P$.

We remind the reader that a ***Steiner triangulation*** of a point set $P$ is a triangulation of $P$ where we are allowed to add additional points.

The crossing number is defined in analogously in higher dimension, where we consider the number of objects intersecting a hyperplane (instead of a line).

Let $P$ be a set of $n$ points in the plane. For two functions $f(P)$ and $g(P)$, we denote by $f(P) \preceq g(P)$ if $f(P) = O(g(P))$ and there exists a set of $n$ points $P$, such that $f(P)$ is smaller than $g(P)$ by a polynomial factor in $n$. Similarly, $f(P) \approx g(P)$ denote the fact that $f(P) = O(g(P)\log^{O(1)} n)$ and $g(P) = O(f(P)\log^{O(1)} n)$.

The relationship between the above quantities is

$$cr_{\mathrm{match}}(P) \preceq cr_{\mathrm{span}}(P) \approx cr_{\triangle}^{St}(P) \preceq cr_{\triangle}(P)\,.$$

For proof of this is provided in Appendix A.

# 3 Approximating the spanning tree with optimal crossing number

Consider a set system $\mathcal{I} = (P, \mathcal{F})$ of finite $\mathsf{VC}$ dimension $\tau$. For more details on spaces with bounded $\mathsf{VC}$ dimension see [PA95]. For our purposes, it is sufficient that $\mathcal{F}$ is a set of subsets of $P$ of cardinality bounded by $|P|^{\tau}$, and this holds for any set system induced by a subset of $P$.

For two distinct points $\mathsf{p}, \mathsf{q} \in P$, we will refer to the set $\{\mathsf{p}, \mathsf{q}\}$ as an ***edge***, denoted by $\mathsf{pq}$. An edge $\mathsf{pq}$ ***crosses*** a set $S \in \mathcal{F}$ if $|\{\mathsf{p}, \mathsf{q}\} \cap S| = 1$.

The ***crossing number*** of a set of edges $F$ of $P$ is the maximum number of edges of $F$ crossed by any set of $\mathcal{F}$.

**Example 3.1.** As a concrete example of such a set system, consider a set $P$ of $n$ points in the plane, and let

$$\mathcal{F} = \left\{ P \cap h^+ \;\middle|\; h^+ \text{ is a halfplane} \right\}.$$

The set system $\mathcal{I} = (P, \mathcal{F})$ in this case has $\mathsf{VC}$ dimension 3, and a spanning tree $T$ of $P$ with crossing number $t$, is a spanning tree of $P$, drawn in the plane by straight segments, such that every line (i.e., the boundary of a halfplane) intersects at most $t$ edges of $T$.

**Lemma 3.2.** *Assume there exists a spanning tree $T$ for $P$ with crossing number $t$. Then, for any subset $X \subseteq P$ there exists a spanning tree with crossing number at most $2t$.*

*Proof*: Convert the spanning tree $T$ of $P$, with crossing number $t$, into a closed cycle $C$ visiting the points of $P$, by doing an Euler tour of $T$, using each edge of $T$ twice. The new cycle $C$ has crossing number $2t$. Next, shortcut the cycle $C$ such that it uses only elements of $X$, by replacing each subpath $\pi_{xy}$ (that uses inner vertices that are not in $X$) connecting $x, y \in X$ by the edge $xy$.

This results in a cycle that visits only the vertices of $X$ and has crossing number $\leq 2t$, as such a shortcutting can only decrease the number of edges crossing a set $S \in \mathcal{F}$. Indeed, consider a subpath $x_1x_2, x_2x_3, \ldots, x_{m-1}x_m$, and observe that if $x_1x_m$ crosses a set $S$, then one of the edges in the path must also cross this set. Thus, replacing a subpath (of a cycle) by an edge reduces the crossing number of the cycle. ■

Let $\mathrm{E}(P)$ denote the set of all edges of $P$, and consider the following LP (parameterized by $t$).

$$\begin{aligned}
\gamma(P,t) = \max & \sum_{\mathsf{pq}\in \mathrm{E}(P)} y_{\mathsf{pq}} && && (1)\\
s.t. & \sum_{\mathsf{pq}\in \mathrm{E}(P), |\mathsf{pq}\cap S|=1} y_{\mathsf{pq}} \leq t && \forall S \in \mathcal{F} \\
& \sum_{\mathsf{q}\in P, \mathsf{q}\neq \mathsf{p}} y_{\mathsf{pq}} \geq 1 && \forall \mathsf{p} \in P && (*)\\
& y_{\mathsf{pq}} \geq 0 && \forall \mathsf{pq} \in \mathrm{E}(P).
\end{aligned}$$

Intuitively, this LP tries to pick as many edges as possible (i.e., $y_{\mathsf{pq}} = 1$ indicates that we pick the edge pq), such that (i) no set is being crossed more than $t$ times, and (ii) every point of $P$ participates in at least one edge that is being picked.

**Remark 3.3.** The above LP might not be feasible, and in such a case $\gamma(P,t)$ is not defined. Naturally, one can modify this LP to first compute the minimal value $t$ for which it is feasible, and then solve the original LP with this value of $t$.

In particular, in this case, we can choose almost any arbitrary target function to optimize the LP for. We had chosen this one since the dual form is convenient to work with, see Section 4.

**Lemma 3.4.** *Consider a set system $\mathcal{I} = (P, \mathcal{F})$ with bounded VC dimension, where $n = |P|$, and let $t$ be a parameter such that $\gamma(P,t)$ is feasible. Then, one can compute (in polynomial time) a set of edges $F$, such that $\mathbf{E}[|F|] = \gamma(P,t)$, and the number of connected components of the graph $(P,F)$ is (in expectation) at most $(9/10)n$. The crossing number of $F$ is $O(t + \log n / \log\log n)$ with high probability.*

*Proof*: We solve the LP (1) and compute $\gamma(P,t)$. Next, for every $\mathsf{pq} \in \mathrm{E}(P)$, if $y_{\mathsf{pq}} \geq 1$ then we add pq to $F$. Otherwise, if $y_{\mathsf{pq}} < 1$ then we pick the edge pq into $F$ with probability $y_{\mathsf{pq}}$.

For a set $S \in \mathcal{F}$, let $X_S$ be its crossing number in $F$. We have that

$$\mu = \mathbf{E}[X_S] \leq \sum_{\substack{\mathsf{pq}\in \mathrm{E}(P),\\ |\mathsf{pq}\cap S|=1}} y_{\mathsf{pq}} \leq t.$$

For a constant $c > 0$ sufficiently large, let $\delta = c + c\,(\log n) \Big/ \left( t \log \dfrac{\log n}{t} \right)$, and observe that $t\delta \log \delta = \Omega(\log n)$. As such, by the Chernoff inequality, we have that

$$\begin{aligned}\mathbf{Pr}[X_S > (1+\delta)t] \leq \mathbf{Pr}\left[X_S > \left(1 + \frac{t\delta}{\mu}\right)\mu\right] &\leq \left(\frac{\exp(t\delta/\mu)}{(1+t\delta/\mu)^{1+t\delta/\mu}}\right)^{\mu} \\ &= \exp\left(t\delta - (\mu + t\delta)\ln\left(1 + \frac{t\delta}{\mu}\right)\right) \leq \exp(t\delta - t\delta\log\delta) < \frac{1}{n^{O(1)}}.\end{aligned}$$

Since $\mathcal{I}$ has $\mathsf{VC}$ dimension $\tau$, the number of sets one has to consider (i.e., the size of $\mathcal{F}$) is $O(n^{\tau})$ [PA95], which implies, by the above, that the crossing number of $F$ is bounded by $(1+\delta)t = O(t + \log n/\log\log n)$, with high probability.

As for the number of connected components in the graph $G = (P, F)$, observe that a point $\mathsf{p}$ is not adjacent to any edge of $F$ with probability

$$\prod_{\substack{\mathsf{q}\in P,\\ \mathsf{q}\neq\mathsf{p}}} (1 - y_{\mathsf{pq}}) \leq \exp\left(-\sum_{\mathsf{q}\in P, \mathsf{q}\neq\mathsf{p}} y_{\mathsf{pq}}\right) \leq \frac{1}{e},$$

by the inequality (*) in $\mathsf{LP}$ (1). Let $Y$ be the number of points of $P$ that are singletons in the graph $(P, F)$. By the above, we have that $\mathbf{E}[Y] \leq n/e$. As such, the expected number of connected components in $(P, F)$ is

$$\leq \mathbf{E}\left[\frac{n - Y}{2} + Y\right] \leq \frac{n}{2} + \frac{n}{e} \leq \frac{9}{10} n.$$

■

**Theorem 3.5.** *Consider a set system $\mathcal{I} = (P, \mathcal{F})$ with bounded $\mathsf{VC}$ dimension, where $n = |P|$. Let $t$ be the minimum crossing number of any spanning tree of $\mathcal{I}$. Then, one can compute, in polynomial time, a spanning tree $T$ of $P$ with a crossing number $O(t \log n + \log^2 n/\log\log n)$.*

*Proof*: We set $P_1 = P$. In the $i$th iteration, we compute the minimal $t_i$ for which $\gamma(P_i, t_i)$ is feasible. Next, we compute a set of edges $F_i$ over $P_i$, using Lemma 3.4. If the number of connected components of $(P_i, F_i)$ is larger than $(19/20)\,|P_i|$, we repeat this iteration (we have constant probability to succeed by Markov's inequality). Next, from each connected component of $P_i$, we pick one point into $P_{i+1}$. We repeat this algorithm till we remain with a single point. This algorithm performs $m = O(\log n)$ iteration. Now, the union $F = \cup_i F_i$ forms a spanning graph of $P$, and we return any spanning tree $T$ of $(P, F)$.

The crossing number of $T$ is bounded by the total crossing numbers of the graphs $G_1 = (P_1, F_1), \ldots, G_m = (P_m, F_m)$. Now, the graph $G_i$ has crossing number $O(t_i + \log n/\log\log n)$ by Lemma 3.4. By Lemma 3.2, $t_i \leq 2t$, for all $i$. As such, the crossing number of $T$ is $O\Big(\sum_i (t_i + \log n/\log\log n)\Big) = O(t\log n + \log^2 n/\log\log n)$. ■

When $P$ is a set of points in $\mathbb{R}^d$, we will be interested in the spanning tree having the minimal number of crossings with any hyperplane. In particular, the above result implies the following.

**Corollary 3.6.** *Let $P$ be a set of $n$ points in $\mathbb{R}^d$. Then, one can compute, in polynomial time, a spanning tree $T$ of $P$ with a crossing number $\Delta = O(cr_{\text{span}}(P)\log n + \log^2 n/\log\log n)$. Specifically, any hyperplane in $\mathbb{R}^d$ crosses at most $\Delta$ edges (i.e., segments) of $T$.*

# 4 Spanning tree in $\mathbb{R}^d$ with low crossing number

Let $P$ be a set of $n$ points in the plane in general position (i.e., no three points are colinear). Let $\mathcal{L}(P)$ denote the set of all partitions of $P$ into two non-empty sets, by a line that does not contain any point of $P$. For each such partition, we select a representative line that realizes this partition. We slightly abuse notations as refers to $\mathcal{L}(P)$ as a set of these lines.

We are interested in the question of finding a spanning tree $T$ of $P$ such that each line of $\mathcal{L}(P)$ crosses at most $O(\sqrt{n})$ edges of $T$.

**Definition 4.1.** For a set of lines $L$ in the plane, the crossing distance between two points is the number of lines of $L$ crossed by the segment formed by these two points. Formally, for any two points $\mathsf{p}, \mathsf{q} \in \mathbb{R}^2$, the ***crossing distance*** between them is $d_L(\mathsf{p}, \mathsf{q}) = x + y/2$, where $x$ is the number of lines of $L$ having $\mathsf{p}$ and $\mathsf{q}$ on opposite sides, and $y$ is the number of lines that contain either $\mathsf{p}$ or $\mathsf{q}$. It is easy to verify that $d_L(\cdot)$ complies with the triangle inequality (as such its a pseudo-metric).

The ***crossing disk*** of radius $r$ centered at a point $\mathsf{p}$, is the set of all vertices of the arrangement $\mathcal{A}(L)$ in crossing distance at most $r$ from $\mathsf{p}$. We denote this "disk" by $\mathsf{D}_L(\mathsf{p}, r)$.

We need the following lemma due to Welzl [Wel92].

**Lemma 4.2 ([Wel92]).** *Let $r \geq 0$ be a parameter, $L$ be a set of lines (of size at least $2r$) in the plane, and let $\mathsf{p}$ be a point in the plane not contained in any line of $L$. Then $|\mathsf{D}_L(\mathsf{p}, r)| \geq \binom{r+1}{2}$.*

Here is the $\mathsf{LP}$ (1) specialized for this planar case, and its dual $\mathsf{LP}$.

$$
\begin{aligned}
\gamma'(P,t) = \max & \sum_{\mathsf{pq}\in \mathrm{E}(P)} x_{\mathsf{pq}} & \\
s.t. \quad & \sum_{\substack{\mathsf{pq}\in \mathrm{E}(P),\\ \mathsf{pq}\cap\ell\neq\emptyset}} x_{\mathsf{pq}} \leq t & \forall \ell \in \mathcal{L}(P)\\
& \sum_{\mathsf{q}\in P, \mathsf{q}\neq \mathsf{p}} x_{\mathsf{pq}} \geq 1 & \forall \mathsf{p} \in P\\
& x_{\mathsf{pq}} \geq 0 & \forall \mathsf{pq} \in \mathrm{E}(P).
\end{aligned}
$$

$$
\begin{aligned}
\alpha'(P,t) = \min \quad & t \sum_{\ell\in\mathcal{L}(P)} z_\ell - \sum_{\mathsf{p}\in P} z_{\mathsf{p}} & \\
s.t. \quad & \sum_{\substack{\ell\in\mathcal{L}(P),\\ \ell\cap\mathsf{pq}\neq\emptyset}} z_\ell - z_{\mathsf{p}} - z_{\mathsf{q}} \geq 1 & \\
& & \forall \mathsf{pq} \in \mathrm{E}(P)\\
& z_\ell \geq 0 & \forall \ell \in \mathcal{L}(P)\\
& z_{\mathsf{p}} \geq 0 & \forall \mathsf{p} \in P.
\end{aligned}
$$

We will next show that $\gamma'(P, \sqrt{n})$ is feasible. This would imply that one can find spanning graph of $P$ with crossing number of $O(\sqrt{n})$ that uses a constant fraction of the vertices (i.e., Lemma 3.4).

**Lemma 4.3.** *The* LP $\gamma'(P,t)$ *is feasible for* $t = \sqrt{n}$.

*Proof*: Consider the dual LP above and observe that it is always feasible (for example by setting $z_\ell = 1$ for all $\ell \in \mathcal{L}(P)$ and $z_\mathsf{p} = 0$ for all $\mathsf{p} \in P$). Thus, if we show that $\alpha'(P,t)$ is bounded from below (and thus is finite), then the strong duality theorem would imply that $\gamma'(P,t)$ is feasible and equal to $\alpha'(P,t)$.

So consider a solution to this dual LP, where all the values are rational numbers. Let $U > 1$ be the smallest integer such that if we scale all the values in the given LP solution by $U$ then they are integers. In particular, let $y_\ell = U z_\ell$, for all $\ell \in \mathcal{L}(P)$, and $y_\mathsf{p} = U z_\mathsf{p}$, for all $\mathsf{p} \in P$.

Let $L$ be a set of lines, where we pick $y_\ell$ copies of $\ell$ into this set, for all $\ell \in \mathcal{L}(P)$. Formally, $\psi$ copies of the same line $\ell$ (put into $L$) will be a collection of $\psi$, almost identical, copies of the line $\ell$ slightly perturbed so that these $\psi$ lines are in general position. Thus, $L$ is a set of $N = U \sum_{\ell\in\mathcal{L}(P)} z_\ell$ lines in general position. Furthermore, the inequalities in the LP implies that, for any segment $\mathsf{pq} \in \mathrm{E}(P)$, we have that $\mathsf{pq}$ crosses

$$d_L(\mathsf{p},\mathsf{q}) = \sum_{\substack{\ell\in\mathcal{L}(P),\\ \ell\cap\mathsf{pq}\neq\emptyset}} y_\ell = U \sum_{\substack{\ell\in\mathcal{L}(P),\\ \ell\cap\mathsf{pq}\neq\emptyset}} z_\ell \geq U\,(1 + z_\mathsf{p} + z_\mathsf{q}) = U + y_\mathsf{p} + y_\mathsf{q}$$

lines of $L$.

Observe that $\mathsf{D}_L(\mathsf{p}, y_\mathsf{p}) \cap \mathsf{D}_L(\mathsf{q}, y_\mathsf{q}) = \emptyset$ for any pair $\mathsf{pq} \in \mathrm{E}(P)$. Otherwise, there would be a point $\mathsf{r}$ in the plane such that $d_L(\mathsf{p},\mathsf{r}) \leq y_\mathsf{p}$ and $d_L(\mathsf{q},\mathsf{r}) \leq y_\mathsf{q}$. But the triangle inequality would imply that $d_L(\mathsf{p},\mathsf{q}) \leq y_\mathsf{p} + y_\mathsf{q}$, which contradicts the above.

By Lemma 4.2, for $\mathsf{p} \in P$, the disk $\mathsf{D}_L(\mathsf{p}, r)$ contains at least $\binom{y_\mathsf{p}+1}{2}$ distinct vertices of $\mathcal{A}(L)$. On the other hand, the total number of vertices in the arrangement $\mathcal{A}(L)$ is $\binom{N}{2}$. We conclude that

$$\frac{U^2}{2}\sum_{\mathsf{p}\in P} z_\mathsf{p}^2 \leq \sum_{\mathsf{p}\in P}\binom{y_\mathsf{p}+1}{2} \leq \binom{N}{2} \leq \frac{U^2}{2}\left(\sum_{\ell\in L} z_\ell\right)^2. \quad\implies\quad \sum_{\mathsf{p}\in P} z_\mathsf{p}^2 \leq \left(\sum_{\ell\in L} z_\ell\right)^2.$$

Now, by the Cauchy-Schwarz inequality and the above, we have that

$$\sum_{\mathsf{p}\in P} z_\mathsf{p} \leq \sqrt{n}\sqrt{\sum_{\mathsf{p}\in P} z_\mathsf{p}^2} \leq \sqrt{n}\sum_{\ell\in L} z_\ell.$$

As $t = \sqrt{n}$ this implies that $\alpha'(n,t) = \sqrt{n}\sum_{\ell\in L} z_\ell - \sum_{\mathsf{p}\in P} z_\mathsf{p} \geq 0$. The claim now follows. ■

**Remark 4.4.** The proof of Lemma 4.3 works also in higher dimensions, where we consider points and hyperplanes in $\mathbb{R}^d$. There, one has to use Hölder's inequality instead of the Cauchy-Schwarz inequality. Then, the LP is feasible for $t = O\big(n^{1-1/d}\big)$.

**Theorem 4.5.** *Given a set $P$ of $n$ points in $\mathbb{R}^d$, one can compute (in polynomial time) a spanning tree $T$ of $P$ with crossing number at most $O\big(n^{1-1/d}\big)$; that is, any hyperplane in $\mathbb{R}^d$ crosses at most $O\big(n^{1-1/d}\big)$ edges of $T$.*

*Proof*: By Lemma 4.3 and Remark 4.4, $\gamma'(P,t)$ is feasible, for $t = O\big(n^{1-1/d}\big)$. As such, by Lemma 3.4, we can compute a set of edges $F$ that engages a constant fraction of the points of $P$, and it has crossing number $O(t + \log n/\log\log n) = O(t)$. Using the algorithm of Theorem 3.5 generates a spanning tree with crossing number $C(n) = O\big(n^{1-1/d}\big) + C((19/20)n)$. Namely, the resulting spanning tree has crossing number $O\big(n^{1-1/d}\big)$. ■

**Remark 4.6 (Connection to separating/hitting and packing LPs.).** It is interesting to consider the LP that just tries to separate all points of $P$ from each other. It looks similar to our dual LP while being simpler.

$$
\begin{array}{lll}
\operatorname{sep}(P) = \min & \displaystyle\sum_{\ell \in \mathcal{L}(P)} z_\ell & \\
s.t. & \displaystyle\sum_{\substack{\ell \in \mathcal{L}(P),\\ \ell \cap \mathsf{pq} \neq \emptyset}} z_\ell \geq 1 & \forall \mathsf{pq} \in \mathrm{E}(P) \\
& z_\ell \geq 0 & \forall \ell \in \mathcal{L}(P) \\
& z_{\mathsf{p}} \geq 0 & \forall \mathsf{p} \in P.
\end{array}
$$

Now, a similar scaling argument to the one used in the proof of Lemma 4.3 implies that $\sum_{\ell \in \mathcal{L}(P)} z_\ell \geq \sqrt{n}/2$. Namely, any fractional set of lines separating $n$ points in the plane is of size $\Omega(\sqrt{n})$. Naturally, this argument works also in higher dimensions, where a fractional set of hyperplanes of size $\Omega(n^{1-1/d})$ is required to separate a set of $n$ points in $\mathbb{R}^d$.

Observe, that since the LP $\alpha'(\gamma,t)$ is more restrictive than this LP, we conclude that for any feasible solution to $\alpha'(\gamma,t)$ it holds that $\sum_{\ell \in \mathcal{L}(P)} z_\ell \geq \sqrt{n}/2$.

The dual to the above LP is the packing LP that tries to pick as many fractional edges as possible, while no line crosses edges with total value exceeding 1. While this is similar to our primal LP $\gamma'(P,t)$, it is not clear how to round it, since we do not have the guarantee that every point has sufficient number of edges attached to it in the fractional solution.

## 4.1 A Deterministic algorithm for the planar case

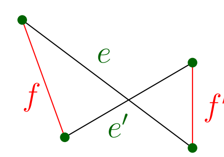
Interestingly, at least in the planar case, one can do the rounding deterministically.

**Lemma 4.7.** *Let $P$ be a set of $n$ points in the plane and a parameter $t$, such that $\gamma'(P,t)$ is feasible. Then, one can compute, in polynomial deterministic time, a set of edges $F$, such that (i) the crossing number of $F$ is $\leq 12t$, and (ii) the number of connected components in $(P, F)$ is $\leq (3/4)n$.*

$$
\begin{array}{llr}
\min \displaystyle\sum_{\mathsf{pq} \in \mathrm{E}(P)} \|\mathsf{p} - \mathsf{q}\|\, x_{\mathsf{pp}} & & (2) \\
s.t. \displaystyle\sum_{\substack{\mathsf{pq} \in \mathrm{E}(P),\\ \mathsf{pq} \cap \ell \neq \emptyset}} x_{\mathsf{pq}} \leq t & \forall \ell \in \mathcal{L}(P) & \\
\displaystyle\sum_{\mathsf{q} \in P, \mathsf{q} \neq \mathsf{p}} x_{\mathsf{pq}} \geq 1 & \forall \mathsf{p} \in P & (*) \\
x_{\mathsf{pq}} \geq 0 & \forall \mathsf{pq} \in \mathrm{E}(P)\,. &
\end{array}
$$

Figure 1: The modified LP

*Proof*: Instead of computing $\rho = \gamma'(P,t)$ we slightly modify the LP so that it finds the "shortest" such solution. The resulting modified LP is depicted in Figure 1.




Let $H$ be the set of all the edges $\mathsf{pq}$ in the solution to this LP such that $x_{\mathsf{pq}} > 0$. We claim that this set of edges is planar. Indeed, if two such segments $e$ and $e'$ intersect, then consider two opposing edges $f$ and $f'$ of the quadrant formed by the convex hull of the endpoints of $e$ and $e'$, see figure on the right.

We have that $\|f\| + \|f'\| < \|e\| + \|e'\|$, which implies that, for $\delta > 0$ sufficiently small, the solution $x_e = x_e - \delta$, $x_{e'} = x_{e'} - \delta$, $x_f = x_f + \delta$, $x_{f'} = x_{f'} + \delta$ is feasible, as the total value of the edges attached to a vertex does not change, and the crossing number of any line does not increase by this change, as can be easily verified. But this implies that there is a feasible solution with a better target value (specifically, the target value goes down by $\delta \cdot (\|e\| + \|e'\| - \|f\| - \|f'\|)$). A contradiction.

Thus $G = (P, H)$ is a planar graph where each point of $P$ has at least one edge attached to it. Furthermore, the average degree in a planar graph is at most 6, which implies that at least half of the points of $P$ have degree at most 12 in $G$. But each such point $\mathsf{p}$, must have an edge $\mathsf{pq}$ attached to it, such that $x_{\mathsf{pq}} \geq 1/12$, because of (*) in the LP (2).

Thus, scale the LP solution by a factor of 12 and pick all the edges $\mathsf{pq}$ with $12x_{\mathsf{pq}} \geq 1$ into the set $F$. We get a set that is adjacent to at least half of the points of $P$, and its crossing number is at most $12t$. ■

The planarity argument in the above proof of Lemma 4.7 is similar to the one used by Fekete *et al.* [FLM08] – they use it to argue that there is one heavy edge, while we use it to argue that there are many heavy edges.

**Theorem 4.8.** *Let $P$ be a set of $n$ points in the plane. One can compute, in deterministic polynomial time, a spanning tree $T$ of $P$ with a crossing number $O\big(\min\big(cr_{\mathrm{span}}(P) \log n, \sqrt{n}\,\big)\big)$.*

# 5 On triangulations with low crossing number

Given a point set $P$ computing the triangulation with the lowest crossing number seems quite challenging, and currently nothing is known about this problem. We make two observations about this problem.

Here, given a triangulation (or a set of triangles) $S$ of $P$, its ***crossing number*** is the maximum number of triangles of $S$ intersected by any line in the plane.

## 5.1 Cover by triangles with low crossing number

In this section, we show how to extract a cover by triangles that has small crossing number.

For a set of points $P$, let $\boxtimes(P)$ denote the set of triangles of $P$ such that they do not contain any other point of $P$. Formally,

$$\boxtimes(P) = \left\{ \triangle\mathsf{pqr} \;\middle|\; \mathsf{p}, \mathsf{q}, \mathsf{r} \in P \text{ and } \triangle\mathsf{pqr} \cap P = \{\mathsf{p}, \mathsf{q}, \mathsf{r}\} \right\}.$$

(We assume for the sake of simplicity of exposition that no three points of $P$ are colinear.) Clearly, the set $\boxtimes(P)$ is of size $O(n^3)$ and it can be computed in $O(n^4)$ time (it can be done even faster, but thats not our worry here).

**Lemma 5.1.** *For a given set $P$ of $n$ points in the plane, let $cr_\triangle(P)$ be the minimum crossing number of any triangulation of $P$. One can compute a set of triangles $S \subseteq \boxtimes(P)$, such that (i) any line $\ell$ in the plane crosses at most $O(cr_\triangle(P) \log n)$ triangles of $S$, (ii) any point inside the convex hull of $P$ is covered by some triangle of $S$, and no point is covered by more than $O(\log n)$ such triangles, and (iii) $|S| = O(n \log n)$.*

*Proof*: Compute the arrangement $\mathcal{A} = \mathcal{A}(\boxtimes(P))$. We write an $\mathsf{LP}$ relaxation that tries to compute the optimal triangulation. To this end, for every face $f \in \mathcal{A}$, we associate an arbitrary point $\mathsf{p}_f \in \mathrm{int}(f)$ that does not lie on any line induced by a pair of points of $P$. Let

$$Q = \left\{ \mathsf{p}_f \;\middle|\; f \in \mathcal{A} \text{ and } \mathsf{p}_f \in \mathcal{CH}(P) \right\}$$

be a set of witness points placed inside each face (for the faces inside the convex hull of $P$), where $\mathcal{CH}(P)$ denotes the convex hull of $P$. Next, we associate a variable $x_\triangle$ with each triangle $\triangle$ of $\boxtimes(P)$ (that decides whether or not we pick it into the triangulation). We now write the cover LP for this problem, requiring that no line has high crossing number.

$$
\begin{aligned}
\psi(P) = \min \quad & t && && (3)\\
s.t. \quad & \sum_{\substack{\triangle \in \boxtimes(P),\\ \triangle \cap \ell \neq \emptyset}} x_\triangle \leq t && \forall \ell \in \mathcal{L}(P) &&\\
& \sum_{\substack{\triangle \in \boxtimes(P),\\ \mathsf{p} \in \triangle}} x_\triangle = 1 && \forall \mathsf{p} \in Q &&\\
& \sum_{\triangle \in \boxtimes(P)} x_\triangle \leq 2n - 4 && && (*)\\
& x_\triangle \geq 0 && \forall \triangle \in \boxtimes(P)\,. &&
\end{aligned}
$$

We solve this LP that has $O(n^3)$ variables, and $O(n^4)$ inequalities. Clearly, the optimal triangulation realizing $cr_\triangle(P)$ is a feasible solution to this LP (as a triangulation of $n$ points has at most $2n-4$ triangles as can be easily verified from the Euler formula). As such $\psi(P) \leq cr_\triangle(P)$. We now pick each triangle $\triangle \in \boxtimes(P)$ into a random sample $S$ with probability $\min(1, cx_\triangle \log n)$, where $c$ is a sufficiently large constant.

By the Chernoff inequality, with high probability, every point of $Q$ is covered by some triangle of $S$. As such, the triangles of $S$ cover the convex hull of $P$ completely. Arguing in a similar fashion, we have that every point of $Q$, with high probability, is covered by at most $O(\log n)$ such triangles of $S$.

Similarly, by the Chernoff inequality, no line of $\mathcal{L}(P)$ crosses more than $O(\psi(P)\log n)$ triangles of $S$. As such, this bound holds for any line in the plane.

Finally, the (*) inequality in LP (3) implies, again by the Chernoff inequality, that $|S| = O(n \log n)$. ■

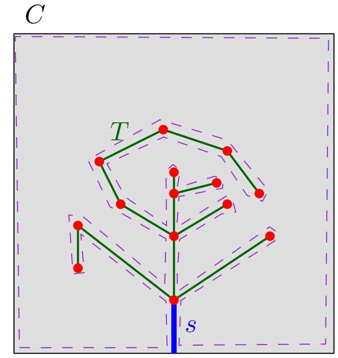
## 5.2 Steiner triangulation with low crossing number

Surprisingly, computing a Steiner triangulation with low crossing number is much easier.

**Lemma 5.2.** *Given a point set $P$, one can compute a Steiner triangulation of $P$ with crossing number $O(cr_{\mathrm{span}}(P)\log^2 n + \log^3 n/\log\log n)$.*

*Proof*: Compute, using Theorem 3.5, a spanning tree $T$ of $P$ with crossing number $t' = O(cr_{\mathrm{span}}(P)\log n + \log^2 n/\log\log n)$.

Next, let $C$ be a slightly expanded axis parallel bounding square of $P$. Consider the region in the plane formed by $\pi = \partial C \cup T \cup s$, where $s$ is the shortest segment connecting $\partial C$ and $T$, see figure on the right. The set $\pi$ is the boundary of a simple polygon with $O(n)$ edges (it is slightly degenerate but it is easy to extend our argument to take care of this minor issue, for example by walking around the boundary, resulting in the dashed polygon in the figure on the right).




Hershberger and Suri [HS95] showed how to Steiner triangulate a simple polygon, with $m$ edges, such that any ray crosses at most $O(\log m)$ edges. Clearly, this triangulation of $\pi$ results in a Steiner triangulation of $P$, covering the square $C$, with crossing number $O(t' \log n)$, as the portion of any line crossing $C$ can be represented as a union of $2t'$ rays (that end when they hit $\pi$), and any such ray crosses $O(\log n)$ triangles. ∎

**Remark 5.3.** Instead of using the bounding square $C$ in the above proof, one can use a slightly enlarged convex hull of $P$. This results with a Steiner triangulation of the convex hull of $P$ with the same guaranteed bound on the crossing number.

**Remark 5.4.** Given a Steiner triangulation of a point set $P$ with crossing number $t$, we can extract a Steiner tree of $P$ with the same crossing number. Next, by doubling edges and shortcutting we get a spanning path of $P$ with crossing number $2t$. As such, we have that $cr_{\mathrm{span}}(P) \leq 2cr^{St}_{\triangle}(P)$.

Combining this remark with Lemma 5.2 implies the following.

**Corollary 5.5.** *Given a point set $P$, one can compute a Steiner triangulation of $P$ with crossing number $O(cr^{St}_{\triangle}(P) \log^2 n + \log^3 n / \log\log n)$, where $cr^{St}_{\triangle}(P)$ is the minimum crossing number of any* Steiner *triangulation of $P$.*

# 6 Conclusions

We presented an approximation algorithm for computing a spanning tree with low crossing number. The new algorithm relies on a natural LP relaxation of the problem and a straightforward rounding scheme.

Interestingly, our approach enables us to provide a direct proof to the existence of such spanning trees in $\mathbb{R}^d$. This is, as far as we know, the first algorithm for this problem that avoids using the reweighting technique. Intuitively, our algorithm (together with previous results [Lon01]) suggests that reweighting in geometric settings can sometimes be replaced by LP rounding. This is a significant feature, as LPs are considerably more general and flexible tool than reweighting. For example, using our algorithm, we can add other constraints to the LP; e.g., we can insist that some certain cuts would have significantly lower crossing number than some other cuts. In particular, it is not clear how one can incorporate such considerations into a reweighting algorithm computing spanning trees with low crossing number.

One interesting open problem, is to compute spanning trees with *relative* crossing number using the new LP approach. Here, given a point set $P$ in $\mathbb{R}^3$, one would like to compute a spanning tree $T$ such that if a halfspace $h^+$ contains $k$ points of $P$ then it boundary plane $h$ crosses (say) $O((k \log n)^{2/3})$ edges of $T$. Such a result is known in the plane [AHS07], but the problem is open in higher dimensions.


## Acknowledgments

The author would like to thank Chandra Chekuri, Jeff Erickson, Sándor Fekete, Jirka Matoušek, and Emo Welzl for helpful discussions on the problems studied in this paper.


## References


[AE98] P. K. Agarwal and J. Erickson. Geometric range searching and its relatives. In B. Chazelle, J. E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*. Amer. Math. Soc., 1998.

[AHK06] S. Arora, E. Hazan, and S. Kale. Multiplicative weights method: a meta-algorithm and its applications. manuscript. Available from `http://www.cs.princeton.edu/~arora/pubs/MWsurvey.pdf`, 2006.

[AHS07] B. Aronov, S. Har-Peled, and M. Sharir. On approximate halfspace range counting and relative epsilon-approximations. In *Proc. 23rd Annu. ACM Sympos. Comput. Geom.*, pages 327–336, 2007.

[BG95] H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:263–279, 1995.

[BGRS04] V. Bilò, V. Goyal, R. Ravi, and M. Singh. On the crossing spanning tree problem. In *8th Intl. Work. Approx. Algs. Combin. Opt. Problems*, pages 51–60, 2004.

[Cla93] K. L. Clarkson. Algorithms for polytope covering and approximation. In *Proc. 3th Workshop Algorithms Data Struct.*, volume 709 of *Lect. Notes in Comp. Sci.*, pages 246–252. Springer-Verlag, 1993.

[CV09] C. Chekuri and J. Vondrák. Dependent randomized rounding for matroid polytopes and applications. *CoRR*, abs/0909.4348, 2009.

[CW89] B. Chazelle and E. Welzl. Quasi-optimal range searching in spaces of finite VC-dimension. *Discrete Comput. Geom.*, 4:467–489, 1989.

[ERS05] G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. *Inform. Process. Lett.*, 95(2):358–362, 2005.

[FLM08] S. P. Fekete, M. E. Lübbecke, and H. Meijer. Minimizing the stabbing number of matchings, trees, and triangulations. *Discrete Comput. Geom.*, 40(4):595–621, 2008.

[HS95] J. Hershberger and S. Suri. A pedestrian approach to ray shooting: shoot a ray, take a walk. *J. Algorithms*, 18(3):403–431, 1995.

[Lon01] P. M. Long. Using the pseudo-dimension to analyze approximation algorithms for integer programming. In *Proc. 7th Workshop Algorithms Data Struct.*, volume 2125 of *Lecture Notes Comput. Sci.*, pages 26–37, 2001.

[Mat92] J. Matoušek. Efficient partition trees. *Discrete Comput. Geom.*, 8:315–334, 1992.

[PA95] J. Pach and P. K. Agarwal. *Combinatorial Geometry.* John Wiley & Sons, 1995.

[PST91] S. A. Plotkin, D. B. Shmoys, and É Tardos. Fast approximation algorithms for fractional packing and covering problems. In *Proc. 32nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 495–504, 1991.

[Wel92] E. Welzl. On spanning trees with low crossing numbers. In *Data Structures and Efficient Algorithms, Final Report on the DFG Special Joint Initiative*, volume 594 of *Lect. Notes in Comp. Sci.*, pages 233–249. Springer-Verlag, 1992.

# A On the relationships between the different crossing numbers

We claim that the following holds

$$cr_{\mathrm{match}}(P) \preceq cr_{\mathrm{span}}(P) \approx cr^{St}_{\triangle}(P) \preceq cr_{\triangle}(P)\,.$$

See Section 2 for relevant definitions.

**Lemma A.1.** *There is a set $P$ of $2n$ points in the plane, such that $cr_{match}(P) = 2$ and $cr_{\mathrm{span}}(P) = \Omega(\sqrt{n})$.*

*As such, we have $cr_{match}(P) \preceq cr_{\mathrm{span}}(P)$.*

*Proof*: Consider the grid $\sqrt{n} \times \sqrt{n}$ and slightly perturb each point such that they are in general position. Let $P'$ denote the resulting point set. It is easy to verify that any spanning tree for $P$ has crossing number $\Omega(\sqrt{n})$.

Next, for every point $\mathsf{p}_i \in P'$ place very close by another point $\mathsf{p}'_i$, such that no line intersect more than two such segments $\mathsf{p}_i\mathsf{p}'_i$, for $i = 1, \ldots, n$. Let $P$ be the resulting set of $2n$

points. Clearly, $P$ has a matching with crossing number 2, but by the above, any spanning tree of $P$ has crossing number $\Omega(\sqrt{n})$.

As for the second claim, given a spanning tree, it is straightforward to extract a matching with similar crossing number by using shortcutting. ∎

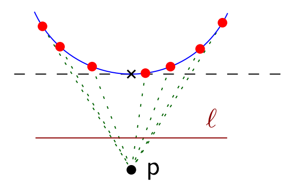
The following is implied by Lemma 5.2 and Remark 5.4.

**Lemma A.2.** *We have* $cr_{\mathrm{span}}(P) \approx cr_{\triangle}^{St}(P)$*.*

**Lemma A.3.** *There is a set of* $n$ *points in the plane, such that* $cr_{\triangle}^{St}(P) = O(\log n)$ *and* $cr_{\triangle}(P) = n - 2$

*As such* $cr_{\triangle}^{St}(P) \preceq cr_{\triangle}(P)$*.*

*Proof*: Let $f(x)$ be a strictly convex non-negative function, such that $f(0) = 0$. Consider a set $P$ of $n-1$ points placed on the image of $f$ for $x$ in the range $[-1, 1]$, and place an extra point at $\mathsf{p} = (0, -2)$. See figure on the right below.




Let $P$ be the resulting set of points. Clearly, any regular triangulation of $P$ requires $\mathsf{p}$ to be connected to all the other points of $P$, and as such a horizontal line $\ell$ slightly above $\mathsf{p}$ will cross $n-2$ triangles of this triangulation.

It is not hard to verify that this point set has a Steiner triangulation with crossing number $O(\log n)$, for example by using the triangulation of Hershberger and Suri [HS95].

The fact that $cr_{\triangle}^{St}(P) \leq cr_{\triangle}(P)$ is obvious. ∎