Robustness of matchings with sub-linear crossing numbers

Mónika Csikós
Université Paris Cité, IRIF, CNRS UMR 8243 and DI ENS, Université PSL.

Nabil H. Mustafa
Université Sorbonne Paris Nord, Laboratoire LIPN, CNRS 7030.

Abstract

Matchings with low crossing numbers were originally introduced in the late 80s for geometric range searching by Welzl [22, 23] and Chazelle-Welzl [5] and have since became a fundamental structure in combinatorics, computational geometry and algorithms.

In this paper, we study the existence of matchings with low crossing numbers with respect to random sampling. In particular, our main theorem states that, given a set system $(X, \mathcal{S})$ with dual VC-dimension d and a parameter $\alpha \in (0, 1]$, a random set of $\tilde{O}\left(n^{1+\alpha}\right)$ edges of $\binom{X}{2}$ contains a linear-sized matching with crossing number $O\left(n^{1-\alpha/d}\right)$. Furthermore, we show that this bound is optimal.

By incorporating the above sampling step to existing algorithms as pre-processing, we immediately obtain improved running times for computing approximate low-crossing matchings for both abstract and geometric set systems.

2012 ACM Subject Classification Theory of computation → Randomness, geometry and discrete structures → Computational geometry

Keywords and phrases Matchings, crossing numbers, approximations

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Funding The work of the authors has been supported by the grants ANR ADDS (ANR-19-CE48-0005) and ANR SAGA (JCJC-14-CE25-0016-01).

1 Introduction

Random sampling is one of the most used and computationally efficient algorithmic tool in computational geometry. Random-sampling based tools with provable guarantees, for instance ε-nets [12, 13] and ε-approximations [21, 14] are of key importance in the field. In this paper, we will focus on uniform sampling for the fundamental problem of computing low-crossing matchings in set systems.

Given a set system $(X, \mathcal{S})$, the *crossing number* of a perfect matching (resp. a spanning tree) G of X with respect to $\mathcal{S}$ is the maximum number of edges of G crossed by any $S \in \mathcal{S}$, where

> a set $S \in \mathcal{S}$ *crosses* a pair $\{x, y\} \subseteq X$ iff $|S \cap \{x, y\}| = 1$.

Matchings with low crossing numbers were originally introduced by Welzl [22, 23] for the special case where X is a set of points in $\mathbb{R}^d$ and $\mathcal{S}$ is induced on X by half-spaces. His result was then improved and generalized by Chazelle and Welzl [5] to a broader class of set systems, using the notion of the dual shatter function $\pi^*_{\mathcal{S}}$ of $(X, \mathcal{S})$:

> For any $k \leq |\mathcal{S}|$, $\pi^*_{\mathcal{S}}(k)$ denotes the maximum number of equivalence classes on X defined by a k-element subfamily $\mathcal{R} \subseteq \mathcal{S}$, where $x, y \in X$ are equivalent with respect to $\mathcal{R}$ if x belongs to the same sets of $\mathcal{R}$ as y.

Their ideas, together with an improvement of one of their claims [11], give the following general theorem.

▶ **Theorem A** ([5, 11]). *Let $(X, \mathcal{S})$ be a set system with $n = |X|$, and dual shatter function $\pi^*_{\mathcal{S}}(k) = O(k^d)$. Then there exists a perfect matching of X with crossing number $O\left(n^{1-1/d} + \ln |\mathcal{S}|\right)$.*

The family of set systems with polynomially bounded dual-shatter function includes set systems with bounded dual VC-dimension and several geometric cases—e.g, the primal and dual set systems induced by half-spaces, balls, and more generally, algebraic inequalities. We refer the reader to the book of Matoušek [16] for details.

Algorithm. The proof of Theorem A is in fact constructive. The original algorithm uses the multiplicative weights update (MWU) technique, as follows. It builds a low-crossing matching iteratively, guided by a weight function ω on $\mathcal{S}$, with initial weights set to 1. At each iteration, the algorithm adds the 'lightest' edge to the matching—that is, the edge that is crossed by sets of minimum total weight. At the end of an iteration, it updates ω by doubling the weight of each set crossing the picked edge—see Algorithm 1.

Algorithm 1 MatchingMWU$\big((X, \mathcal{S})\big)$

```
ω_1(S) ← 1 for all S ∈ 𝒮
X_1 ← X
for i = 1, ..., n/2 do
    e_i ← the lightest edge in (X_i choose 2) w.r.t. ω_i
    Obtain ω_{i+1} from ω_i by doubling the weights of each set crossing e_i
    X_{i+1} ← X_i \ endpoints(e_i)
return {e_1, ..., e_{n/2}}
```

At each iteration, the most expensive step is finding the edge which crosses sets of minimum total weight. Using the incidence matrix for $\mathcal{S}$, this can be done in $O(|E| \cdot |\mathcal{S}|)$ steps, and thus the algorithm has overall time complexity $O(|X| \cdot |E| \cdot |\mathcal{S}|)$, where E is the set of all edges spanned by X. In the past three decades, there several alternative algorithmic approaches were proposed for abstract and geometric set systems; e.g., see [8, 10, 3, 9, 1, 2, 6].

Interestingly, random sampling via ε-nets has been used—both to prove a precise bound, as well as for computational efficiency reasons—for various improvements on Theorem A. In particular, for certain geometric instances such as those induced by half-spaces, it has been shown that a uniform random sample $\mathcal{S}' \subseteq \mathcal{S}$ of size $\tilde{O}(|X|)$ is a good 'test-set'—that is, it suffices to compute a low-crossing matching of X with respect to the sets of $\mathcal{S}'$ only [15]. This result immedately gives an improved runtime; we refer the reader to the book of Chazelle [4] for more details and proofs.

Our Result

The goal of this paper is to study the construction and existence of matchings on a uniform random subset of edges of $\binom{X}{2}$. Surprisingly, as we will later prove as our main result, there exists a matching with low crossing number on a sparse uniform random subset of edges—more precisely, we show that with high probability,

> a random set of $\tilde{O}\left(n^{1+\alpha}\right)$ edges contains a linear-sized matching with crossing number $O\left(n^{1-\alpha/d}\right)$.

Note that we recover the classical bound by setting $\alpha = 1$! As a side-product, one can deduce that there are exponentially many matchings with a low crossing number.

We find this surprising for the following reason: the classical proof of existence (based on the MWU method) assigns exponentially-increasing weights to the sets of $\mathcal{S}$, which then dictate the choice of the edge picked at each iteration. Thus a different choice of the edge at iteration i could result in a changing of the weight distribution, which then influences the sequence of edges picked for all later iterations. At first glance, a random sample chosen *once*, and *uniformly* from $\binom{X}{2}$ cannot simply assure that it will contain many edges from all possible paths chosen by the algorithm.

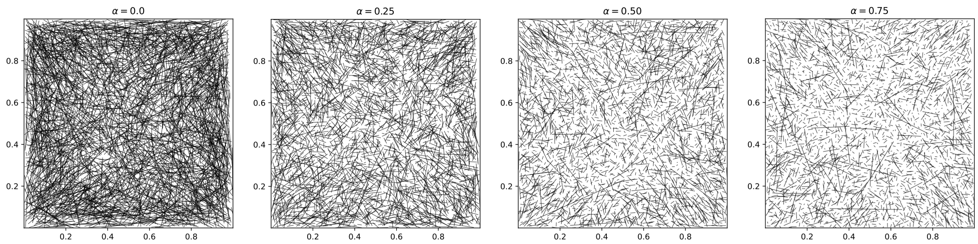
This scenario constrasts sharply with the proof and analysis of other random sampling notions such as ε-nets or ε-approximations which capture 'static' structures. Indeed, in our case, the analysis is subtle: if we fix an initial uniform sample of edges, and build the matching using this sample (by always choosing a light edge from the sample, as in previous algorithms), it introduces a bias (as the set of uncovered points depend on the initial sample of edges) and we cannot assume anymore that among the uncovered points, every edges is picked i.i.d. with a fixed probability. Indeed, with later iterations, the possible paths to be taken care of increase exponentially and the initial random sample does poorly if we use the original analysis.

Therefore, we take a different approach in the statement and the analysis:

1. instead of a perfect matching, we aim for a *linear-sized* partial matching with crossing number $O\left(|X|^{1-\alpha/d}\right)$;
2. instead of the classical algorithm, we propose a new randomized version where an edge from our random sample of edges is picked in each iteration with a carefully chosen probability distribution that *does* depend on the changing weights at each iteration.

We show that these two compromises are sufficient to show an upper bound on the sample size, while also allowing, via $\log |X|$ recursive calls, to be able to construct an approximately low-crossing perfect matching with crossing number $O\left(|X|^{1-\alpha/d}\right)$.

> To visually illustrate the impact of restricting the matching to only use edges of a uniform sample of size $O\left(|X|^{1+\alpha}\right)$, we perfomed some experiments on set systems induced by half-spaces in two dimensions. To do this, we added the pre-sampling feature to an existing implementation [6]. Our input point-sets consist of $n = 5000$ points picked randomly, one from each cell of a uniform grid of size. The matchings with varying α values conforms closely to our theoretical bounds:



We conclude by stating our main theorem:

▶ **Theorem 1** (Main theorem). *Let $(X, \mathcal{S})$ be a set system with dual shatter function $\pi^*(k) = O(k^d)$, and let $\alpha \in (0, 1]$, $\delta \in (0, 1)$ be two given parameters. Let $n = |X|$ and E be a uniform random sample from $\binom{X}{2}$, where each edge is picked i.i.d. with probability*

$$p = \min\left\{\frac{2\ln n}{n^{1-\alpha}} + \frac{4\ln(2/\delta)}{n^{2-\alpha}},\ 1\right\}.$$

Then with probability at least $1-\delta$, E contains a matching of size $n/4$, with crossing number

$$O\left(n^{1-\alpha/d} + \ln |\mathcal{S}|\right).$$

Moreover, we show that the above bound is near-optimal.

▶ **Theorem 2.** *For any $d \geq 2$ and $n_0 \in \mathbb{N}$, there is a set system $(X, \mathcal{S})$ with $|X| = n \geq n_0$, and dual shatter function $\pi^*_{\mathcal{S}}(k) = O(k^d)$, such that the following holds:*

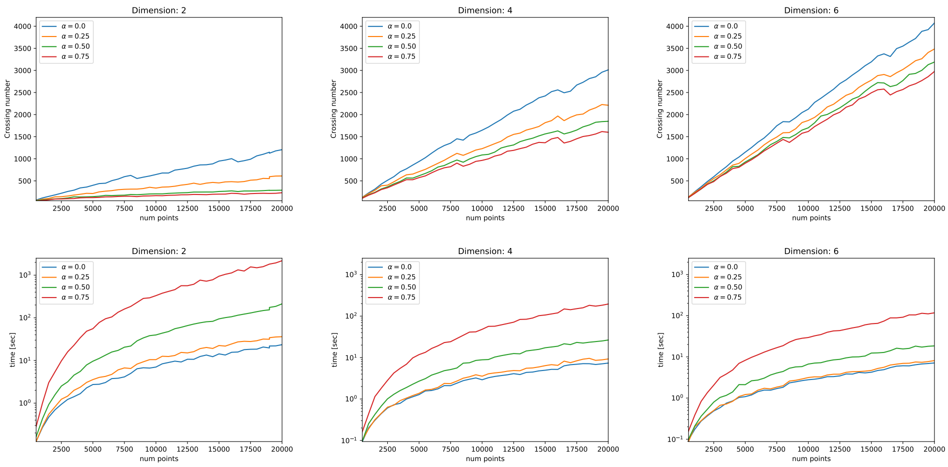
> *let E be a random edge-set obtained by selecting each edge in $\binom{X}{2}$ i.i.d. with probability $p(n) = o\left(n^{\alpha-1}\right)$. Then with probability at least $1/2$, every matching in E of size $n/4$ has crossing number $\omega\left(n^{1-\alpha/d}\right)$ with respect to $\mathcal{S}$.*

Algorithmic consequences: sampling as preprocessing

While Theorem 1 only guarantees the existence of a low-crossing *partial* matching, it can be used to speed up existing algorithms. The only modification we need to make is to build only a matching covering $|X|/16$ elements and then recurse on the uncovered points (note that each recursive subproblem takes a new random sample of edges).
By combining Theorem 1 with the algorithm of Csikós and Mustafa [6], we obtain the new fastest algorithm for constructing matchings with sub-linear crossing numbers, improving previous bound by nearly a factor of n (proof left for full version):

▶ **Corollary 3.** *Let* $(X, \mathcal{S})$ *be a set system with dual shatter function* $\pi^*_{\mathcal{S}}(k) = O(k^d)$*,* $n = |X|$ *and* $\alpha \in (0, 1]$*. Then there is a randomized algorithm which returns a matching of expected crossing number* $O\left(n^{1-\alpha/d} + \ln |\mathcal{S}| \ln n\right)$ *in time* $O\left(n^{1+\alpha+2/d} + |\mathcal{S}| \cdot n^{2/d}\right)$*.*

The figures below show the asymptotic behavior of the guarantees and the running times for the modification of [6]. We see that as the dimension increases, the crossing numbers (in linear scale) get closer, while the gap in the running times (in log scale) remains significant.



Before we move on to the proofs of Theorem 1 and other results, we make two remarks.

Remark 1: The original method of Csikós and Mustafa providing matchings with crossing number $O\left(|X|^{1-1/d}\right)$ had running time $O\left(|X|^{2+2/d} + |\mathcal{S}| \cdot |X|^{2/d}\right)$. While $|\mathcal{S}|$ can be significantly larger than $|X|$, Corollary 3 gives improved running times to set systems where succinct test sets exist. Such an instance is geometric set systems induced by half-spaces. In this case, it is sufficient to work with a set of $\tilde{O}(|X|)$ half-spaces that well-represent all half-space ranges [15].

Remark 2: This result implies interesting variants of algorithmic applications of matchings or spanning paths with sub-linear crossing numbers, for instance, in discrepancy theory [17] or algorithmic graph theory [7].

Remark 3: We illustrate the versatility of Theorem 1 by applying it to the MWU algorithm of Welzl and Chazelle-Welzl, to derive a new, more efficient alternative (see Algorithm 2). The following corollary analyses this algorithm; for completeness, we outline its analysis in Section 4.

▶ **Corollary 4.** *Let* $(X, \mathcal{S})$ *be a set system with dual shatter function* $\pi^*_{\mathcal{S}}(k) = O(k^d)$*,* $n = |X|$ *and* $\alpha \in (0, 1]$*. Then algorithm* RecursiveMWU$\big((X, \mathcal{S}), \alpha\big)$ *returns a matching of expected crossing number* $O\left(n^{1-\alpha/d} + \ln |\mathcal{S}| \ln n\right)$ *in time* $O\left(|\mathcal{S}| \cdot n^{2+\alpha}\right)$*.*

Algorithm 2 RecursiveMWU$\big((X, \mathcal{S}), \alpha\big)$

```
M ← ∅
while |X| > 4 do
    ω_1(S) ← 1 for all S ∈ 𝒮
    n ← |X|
    E ← sample of O(n^{1+α} ln n) edges from (X choose 2)
    for i = 1, ..., n/16 do
        e_i ← the lightest edges in E ∩ (X_i choose 2) w.r.t. ω_i
        Define ω_{i+1} from ω_i by doubling the weights of each set in 𝒮
          crossing e_i
        X ← X \ endpoints(e_i)
        M ← M ∪ {e_i}
Match the remaining points of X arbitrarily and add the edges to M
return M
```

2 Proof of Main Theorem (Theorem 1)

The proof is based on the analysis of the following randomized algorithm.

Algorithm 3 RelaxedMWU$\big((X, \mathcal{S}), \alpha, E\big)$

```
ω_1(S) ← 1 for all S ∈ 𝒮
X_1 ← X
for i = 1, ..., n/2 do
    ℰ_i ← the |X_i|^{2−α} lightest edges in (X_i choose 2) w.r.t. ω_i
    if E ∩ ℰ_i = ∅ then
        set T = i − 1 and return {e_1, ..., e_{i−1}}
    else
        Pick an edge e_i from E ∩ ℰ_i uniformly at random
        Define ω_{i+1} from ω_i by doubling the weights of each set
          crossing e_i
        X_{i+1} ← X_i \ endpoints(e_i)
set T = n/2 and return {e_1, ..., e_{n/2}}
```

In particular, Theorem 1 is implied by the following two properties of RelaxedMWU:

▶ **Lemma 5.** *For any halting time $T = t$, the edges returned by* RelaxedMWU *have crossing number $O\left(t^{1-\alpha/d}\right)$.*

▶ **Lemma 6.** *If $E \subseteq \binom{X}{2}$ is an i.i.d. sample where each edge is picked with probability*

$$p = \min\left\{\frac{2\ln n}{n^{1-\alpha}} + \frac{4\ln(2/\delta)}{n^{2-\alpha}},\ 1\right\},$$

then $T \geq n/4$ with probability at least $1-\delta$. In other words, RelaxedMWU$((X, \mathcal{S}), \alpha, E)$ *returns at least $n/4$ edges with probability at least $1-\delta$.*

We continue by proving these two properties of RelaxedMWU separately.

Lemma 5: Bounding the crossing number of the output

Proof of Lemma 5. Let $\{e_1, \ldots, e_t\}$ be the output of RelaxedMWU and for each $i \in [1, t]$, let η_i denote the total weight, **w.r.t.** $\boldsymbol{\omega_i}$, of the sets that cross e_i:

$$\eta_i = \sum_{S \text{ crosses } e_i} \omega_i(S).$$

Furthermore, for any function $f : \mathcal{S} \to \mathbb{R}$, define $f(\mathcal{S}) := \sum_{S \in \mathcal{S}} f(S)$. We will use the following key lemma which gives an upper-bound on η_i (its proof will be presented later).

▶ **Lemma 7.** *Let $(X, \mathcal{S})$ be a set system with dual shatter function $\pi^*_{\mathcal{S}}(k) \le c_1 \cdot k^d$. Then given any $Y \subset X$, weight function $w : \mathcal{S} \to \mathbb{Z}^+$, and $\ell \in \left[|Y|, \binom{|Y|}{2}\right]$, there are at least ℓ distinct edges in $\binom{Y}{2}$ such that each of these edges is crossed by sets of $\mathcal{S}$ of total weight at most*

$$(10c_1)^{1/d} \cdot \frac{w(\mathcal{S}) \cdot \ell^{1/d}}{|Y|^{2/d}}.$$

At the start of iteration i, we have $|X_i| = n - 2i + 2$ and we pick one of the $|\mathcal{E}_i| = |X_i|^{2-\alpha}$ lightest edges of $\binom{X_i}{2}$ w.r.t. ω_i. Applying Lemma 7 with $Y = X_i$, $\omega = \omega_i$ and $\ell = |\mathcal{E}_i|$, we get that each edge in $\mathcal{E}_i$ is crossed by sets of total weight at most

$$(10c_1)^{1/d} \cdot \frac{\omega_i(\mathcal{S}) \cdot |X_i|^{(2-\alpha)/d}}{|X_i|^{2/d}} = (10c_1)^{1/d} \cdot \frac{\omega_i(\mathcal{S})}{|X_i|^{\alpha/d}} = \frac{(10c_1)^{1/d}\omega_i(\mathcal{S})}{(n-2i+2)^{\alpha/d}}, \tag{1}$$

which upper-bounds η_i (since $e_i \in \mathcal{E}_i$).

Let κ_t denote the maximum number of edges in $\{e_1, \ldots, e_t\}$ that are crossed by a set in $\mathcal{S}$. By the weight-update rule of the algorithm, we have

$$\omega_{t+1}(\mathcal{S}) \ge \max_{S \in \mathcal{S}} \omega_{t+1}(S) = 2^{\kappa_t}.$$

On the other hand, we have

$$\omega_{t+1}(\mathcal{S}) = \omega_t(\mathcal{S})\left(1 + \frac{\eta_t}{\omega_t(\mathcal{S})}\right) = \cdots = |\mathcal{S}| \cdot \prod_{j=1}^{t}\left(1 + \frac{\eta_j}{\omega_j(\mathcal{S})}\right) \le |\mathcal{S}| \cdot \exp\left(\sum_{j=1}^{t} \frac{\eta_j}{\omega_j(\mathcal{S})}\right).$$

Combining the upper and lower bounds for $\omega_{t+1}(\mathcal{S})$, we get

$$2^{\kappa_t} \le \omega_{t+1}(\mathcal{S}) \le |\mathcal{S}| \cdot \exp\left(\sum_{j=1}^{t} \frac{\eta_j}{\omega_j(\mathcal{S})}\right) \quad \Longrightarrow \quad \kappa_t \le \frac{1}{\ln 2}\left(\ln|\mathcal{S}| + \sum_{j=1}^{t} \frac{\eta_j}{\omega_j(\mathcal{S})}\right). \tag{2}$$

Using the upper-bound on η_j from Equation (1), we conclude that for any stopping time $t \in [1, n/2]$, the matching $\{e_1, \ldots, e_t\}$ returned by RelaxedMWU has crossing number at most

$$\frac{\ln|\mathcal{S}|}{\ln 2} + \frac{(10c_1)^{1/d}}{\ln 2}\sum_{j=1}^{t} \frac{1}{(n-2j+2)^{\alpha/d}} \le \frac{\ln|\mathcal{S}|}{\ln 2} + \frac{(10c_1)^{1/d}}{\ln 2} \cdot \frac{t^{1-\alpha/d}}{1-\alpha/d} = O\left(t^{1-\alpha/d} + \ln|\mathcal{S}|\right). \tag{3}$$

◀

We will need the following two results for the proof of Lemma 7.

▶ **Theorem 8** (Turán's Theorem [20]). *Let $G = (V, E)$ be a graph with no clique of size $r + 1$. Then*

$$|E| \le \left(1 - \frac{1}{r}\right) \frac{n^2}{2}.$$

▶ **Lemma 9** (Packing Lemma [11]). *Let $(X, \mathcal{S})$ be a set system with shatter function $\pi_{\mathcal{S}}(k) \le c_1 \cdot k^d$ and let $\delta \in (1, |X|)$ be a parameter. Furthermore, let $\mathcal{P} \subset \mathcal{S}$ be a δ-separated set; that is, $|S_1 \Delta S_2| \ge \delta$ for all $S_1, S_2 \in \mathcal{P}$. Then*

$$|\mathcal{P}| \le 2c_1 \left(\frac{|X|}{\delta}\right)^d.$$

Proof of Lemma 7. Let $(\mathcal{S}_w, \mathcal{R}_Y)$ denote the set system where $\mathcal{S}_w$ contains $w(S)$ copies of each $S \in \mathcal{S}$, and

$$\mathcal{R}_Y = \{R_y : y \in Y\}, \quad \text{where} \quad R_y = \{S \in \mathcal{S}_w : y \in S\}.$$

Observe that

- for any $x, y \in Y$, the set $R_x \Delta R_y$ contains precisely the sets in $\mathcal{S}_w$ that cross the edge xy,
- $|\mathcal{S}_w| = w(\mathcal{S})$, and
- the shatter function of $(\mathcal{S}_w, \mathcal{R}_Y)$ is the dual shatter function of $(Y, \mathcal{S}|_Y)$.

Consider a graph G on Y, where there is an edge between two elements $x, y \in Y$ in G if and only if xy is crossed by greater than δ_ℓ sets in $\mathcal{S}_w$, where

$$\delta_\ell = \left(10c_1 \cdot \frac{w(\mathcal{S})^d \ell}{|Y|^2}\right)^{1/d}.$$

Now the **Packing Lemma** implies that any δ_ℓ-separated subset of sets in $\mathcal{R}_Y$ has cardinality at most

$$C_\ell = 2c_1 \left(\frac{w(\mathcal{S})}{\delta_\ell}\right)^d = 2c_1 \frac{w(\mathcal{S})^d}{10c_1 \cdot \frac{w(\mathcal{S})^d \ell}{|Y|^2}} = \frac{|Y|^2}{5\ell}.$$

This implies that G does not contain a clique on $C_\ell + 1$ vertices, and so by **Turán's Theorem**, the number of pairs that are *not* edges in G is at least

$$\binom{|Y|}{2} - \left(1 - \frac{1}{C_l}\right) \frac{|Y|^2}{2} = \frac{|Y|^2}{2C_\ell} - \frac{|Y|}{2} = \frac{5\ell}{2} - \frac{|Y|}{2} \ge \ell,$$

where we used that $|Y| \le \ell$. That is, there are at least ℓ edges which cross sets of total weight at most δ_ℓ. This concludes the proof of Lemma 7. ◀

Lemma 6: Lower-bounding the matching size.

Proof of Lemma 6. Our goal is to prove that if $E \subseteq \binom{X}{2}$ is a random set of edges, where each edge from $\binom{X}{2}$ is picked i.i.d. with probability

$$p = \min\left\{\frac{2\ln n}{n^{1-\alpha}} + \frac{4\ln(2/\delta)}{n^{2-\alpha}},\ 1\right\},$$

then

$$\mathbb{P}[\, T \le n/4 \,] \le \delta.$$

Recall that we have two different sources of randomness: we run the algorithm on an initial sample E of edges and at each iteration, if $E \cap \mathcal{E}_i \neq \emptyset$, we sample an edge $e_i \in E \cap \mathcal{E}_i$ uniformly at random. The notation $\mathbb{P}[A]$ denotes the probability of event A under both sources randomness.

If $p = 1$, then the statement is trivially true, therefore assume that $p < 1$.

We will upper bound the probabilities $\mathbb{P}[\,T = i\,]$ for each $i = 0, \ldots, n/4$. Since E is an i.i.d. uniform random sample of $\binom{X}{2}$,

$$\mathbb{P}[\,T = 0\,] = \mathbb{P}[\,E \cap \mathcal{E}_1 = \emptyset\,] = (1-p)^{|\mathcal{E}_1|} = (1-p)^{n^{2-\alpha}}.$$

Now consider the main case $i \geq 1$. Observe that the edge-set $\mathcal{E}_i$ depends on the edges chosen in earlier iterations $j < i$. To signify this,

for any sequence of $i-1$ edges $(e^1, \ldots, e^{i-1})$, let $\mathcal{E}_i(e^1, \ldots, e^{i-1})$ denote the set of $(n - 2(i-1))^{2-\alpha}$ lightest edges *assuming* that $e^1, \ldots, e^{i-1}$ are added to the matching and the weights of the sets of $\mathcal{S}$ adjusted multiplicatively accordingly.
We say that a sequence $(e^1, \ldots, e^i)$ is *feasible* if

$$e^1 \in \mathcal{E}_1,\ e^2 \in \mathcal{E}_2(e^1),\ \ldots,\ e^i \in \mathcal{E}_i(e^1, \ldots, e^{i-1}).$$

We write $\mathbf{c}^i = (e^1, \ldots, e^i)$ for a sequence of edges representing possible choices of the algorithm over the i iterations, with the convention that $\mathbf{c}^0 = \emptyset$, and $\mathbf{c}^j = (e^1, \ldots, e^j)$.

Let $\mathbf{e}_i = (e_1, \ldots, e_i)$ denote the sequence of random variables representing the edges chosen at each step by the algorithm, with $\mathbf{e}_0 = \emptyset$ and $\mathbf{e}_j = (e_1, \ldots, e_j)$.

We break the analysis into three steps.

1. Unfolding the probability $\mathbb{P}[\,T = i\,]$.

Given the above notation, we have

$$\begin{aligned}
&\mathbb{P}[\,T = i\,] = \mathbb{P}[\,E \cap \mathcal{E}_1 \neq \emptyset,\ \ldots,\ E \cap \mathcal{E}_i \neq \emptyset,\ E \cap \mathcal{E}_{i+1} = \emptyset\,] \\
&= \sum_{\mathbf{c}^i \text{ feasible}} \mathbb{P}\left[\,E \cap \mathcal{E}_1 \neq \emptyset,\ E \cap \mathcal{E}_2(\mathbf{c}^1) \neq \emptyset,\ \ldots,\ E \cap \mathcal{E}_i(\mathbf{c}^{i-1}) \neq \emptyset,\ E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset \;\middle|\; \mathbf{e}_i = \mathbf{c}^i\,\right] \\
&\qquad\qquad \cdot \mathbb{P}\left[\,\mathbf{e}_i = \mathbf{c}^i\,\right] \\
&\leq \sum_{\mathbf{c}^i \text{ feasible}} \mathbb{P}\left[\,E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset \;\middle|\; \mathbf{e}_i = \mathbf{c}^i\,\right] \cdot \mathbb{P}\left[\,\mathbf{e}_i = \mathbf{c}^i\,\right].
\end{aligned}$$

Note that $\mathcal{E}_{i+1}(\mathbf{c}^i)$ is a fixed set once we are given $\mathbf{c}^i = (e^1, \ldots, e^i)$. Using Bayes' theorem, we can express the conditional probabilities on the R.H.S. of the above inequality as

$$\mathbb{P}\left[\,E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset \;\middle|\; \mathbf{e}_i = \mathbf{c}^i\,\right] = \frac{\mathbb{P}\left[\,\mathbf{e}_i = \mathbf{c}^i \;\middle|\; E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\,\right] \cdot \mathbb{P}\left[\,E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\,\right]}{\mathbb{P}[\,\mathbf{e}_i = \mathbf{c}^i\,]}.$$

Thus, we get

$$\mathbb{P}[\,T = i\,] \leq \sum_{\mathbf{c}^i \text{ feasible}} \mathbb{P}\left[\,\mathbf{e}_i = \mathbf{c}^i \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\,\right] \cdot \mathbb{P}\left[\,E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\,\right]$$

since in each term, $\mathbf{c}^i$ is a fixed set of edges and thus $\mathcal{E}_{i+1}(\mathbf{c}^i)$ is not random, we have

$$= \sum_{\mathbf{c}^i \text{ feasible}} \mathbb{P}\left[\ \mathbf{e}_i = \mathbf{c}^i \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\ \right] \cdot (1-p)^{|\mathcal{E}_{i+1}(\mathbf{c}^i)|}.$$

We now proceed by bounding the probability $\mathbb{P}\left[\ \mathbf{e}_i = \mathbf{c}^i \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\ \right]$, iteration by iteration:

$$\begin{aligned}
&\mathbb{P}\left[\ \mathbf{e}_i = \mathbf{c}^i \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\ \right] \\
&= \mathbb{P}\left[\ e_i = e^i \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset,\ \mathbf{e}_{i-1} = \mathbf{c}^{i-1}\ \right] \cdot \mathbb{P}\left[\ \mathbf{e}_{i-1} = \mathbf{c}^{i-1} \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset\ \right] \\
&\ \ \vdots \\
&= \prod_{j=1}^{i} \mathbb{P}\left[\ e_j = e^j \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset, \mathbf{e}_{j-1} = \mathbf{c}^{j-1}\ \right],
\end{aligned}$$

recalling that $\mathbf{e}_0 = \mathbf{c}^0 = \emptyset$. In conclusion,

$$\mathbb{P}\left[\ T = i\ \right] \leq \sum_{\mathbf{c}^i \text{ feasible}} (1-p)^{|\mathcal{E}_{i+1}(\mathbf{c}^i)|} \cdot \prod_{j=1}^{i} \mathbb{P}\left[\ e_j = e^j \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset, \mathbf{e}_{j-1} = \mathbf{c}^{j-1}\ \right]. \quad (4)$$

2. Bounding the probabilities $\mathbb{P}\left[\ e_j = e^j \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset,\ \mathbf{e}_{j-1} = \mathbf{c}^{j-1}\ \right],\ j \in [1, i]$.

Note that, given the conditioning $\mathbf{e}_{j-1} = \mathbf{c}^{j-1}$, and the fact that e_j was picked uniformly from $E \cap \mathcal{E}_j(\mathbf{c}^{j-1})$, we have

$$\begin{aligned}
&\mathbb{P}\left[\ e_j = e^j \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset,\ \mathbf{e}_{j-1} = \mathbf{c}^{j-1}\ \right] \\
&= \sum_{\substack{S' \subseteq \mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i): \\ e^j \in S'}} \mathbb{P}\left[\ e_j = e^j \;\middle|\; E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset,\ \mathbf{e}_{j-1} = \mathbf{c}^{j-1},\ E \cap \mathcal{E}_j(\mathbf{c}^{j-1}) = S'\ \right] \\
&\qquad\qquad \cdot \mathbb{P}\left[\ E \cap \mathcal{E}_j(\mathbf{c}^{j-1}) = S' \;\middle|\; E \cap \mathcal{E}_{i+1}(\mathbf{c}^i) = \emptyset,\ \mathbf{e}_{j-1} = \mathbf{c}^{j-1}\ \right] \\
&= \sum_{\substack{S' \subseteq \mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i): \\ e^j \in S'}} \frac{1}{|S'|} \cdot p^{|S'|} \cdot (1-p)^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)| - |S'|}
\end{aligned}$$

Now rearranging the sum by the size of S', over all S' containing e^j:

$$\begin{aligned}
&= \sum_{\ell=1}^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \binom{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)| - 1}{\ell - 1} \cdot \frac{1}{\ell} \cdot p^{\ell} \cdot (1-p)^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)| - \ell} \\
&= \sum_{\ell=1}^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \frac{1}{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \binom{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|}{\ell} \\
&\qquad\qquad \cdot p^{\ell} \cdot (1-p)^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)| - \ell}
\end{aligned}$$

Using the Binomial theorem, and adjusting for the case $\ell = 0$,

$$\begin{aligned}
&= \frac{1}{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \left((p + (1-p))^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} - (1-p)^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \right) \\
&= \frac{1}{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \left(1 - (1-p)^{|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|} \right).
\end{aligned}$$

In conclusion, we get that $\mathbb{P}\left[\ e_j = e^j \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i)\ \right]$ can be expressed in the form

$$\mathbb{P}\left[\ e_j = e^j \mid E \cap \mathcal{E}_{i+1}(\mathbf{c}^i)\ \right] = \frac{1}{a}\left(1 - (1-p)^a\right) \le p, \tag{5}$$

where $a = |\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)|$ and the last bound follows from $|\mathcal{E}_j(\mathbf{c}^{j-1}) \setminus \mathcal{E}_{i+1}(\mathbf{c}^i)| \ge 1$.

3. Putting everything together.

Let $k_i = \left|\mathcal{E}_i(\mathbf{c}^{i-1})\right|$ and recall that $\left|\mathcal{E}_i(\mathbf{c}^{i-1})\right| = |X_i|^{2-\alpha} = (n - 2(i-1))^{2-\alpha}$.
Continuing Equation (4) together with the bound from Equation (5), we get

$$\mathbb{P}\left[\ T = i\ \right] \le (1-p)^{k_{i+1}} \cdot \sum_{\mathbf{c}^i \text{ feasible}} \prod_{j=1}^{i} p = (1-p)^{k_{i+1}} \cdot k_1 \cdot k_2 \cdots k_i \cdot p^i.$$

Summing the above over all iterations, and using that $k_1 \ge k_2 \ge \cdots \ge k_i$, we get

$$\mathbb{P}\left[\ T \le i\ \right] \le (1-p)^{k_1} + \sum_{\ell=1}^{i} (1-p)^{k_{\ell+1}} \cdot k_1 \cdot k_2 \cdots k_\ell \cdot p^\ell \le (1-p)^{k_{i+1}} + \sum_{\ell=0}^{i} (k_1 \cdot p)^\ell .$$

Thus, for $i = n/4$, we obtain

$$\begin{aligned}\mathbb{P}\left[\ T \le n/4\ \right] &\le (1-p)^{k_{n/4+1}} \sum_{\ell=0}^{n/4} (k_1 \cdot p)^\ell = (1-p)^{k_{n/4+1}} \frac{1 - (pk_1)^{n/4+1}}{1 - pk_1} \\ &\le 2(1-p)^{k_{n/4+1}} \cdot (pk_1)^{n/4} \le 2\exp(-pk_{n/4+1}) \cdot k_1^{n/4}.\end{aligned}$$

Substituting $k_1 = n^{2-\alpha}$, $k_{n/4+1} = (n/2)^{2-\alpha} \ge n^{2-\alpha}/4$ and $p = \frac{2\ln n}{n^{1-\alpha}} + \frac{4\ln(2/\delta)}{n^{2-\alpha}}$, we conclude

$$\mathbb{P}\left[\ T \le n/4\ \right] \le 2\exp\left(-\frac{n \ln n}{2} - \ln\frac{2}{\delta}\right) \cdot \left(n^{2-\alpha}\right)^{n/4} = 2 \cdot \frac{n^{n/2 - \alpha n/4}}{n^{n/2}} \cdot \frac{\delta}{2} \le \delta.$$

Therefore, with probability at least $1 - \delta$, RelaxedMWU returns a matching of size $n/4$. This, together with Lemma 5 concludes the proof of Theorem 1.

◀

3 Proof of Theorem 2

Proof. Now we present the lower-bound construction that shows the optimality of Theorem 1: it is a geometric set system induced by half-spaces on a subset of the integer grid. More precisely, let X be the set of $n = \left\lceil n_0^{1/d} \right\rceil^d$ points defined as $\times_{i=1}^{d} \left[\ 1, \left\lceil n_0^{1/d} \right\rceil\ \right] \subset \mathbb{Z}^d$ and let $\mathcal{S}$ consist of the $d \cdot \left\lfloor n_0^{1/d} \right\rfloor$ subsets of X induced by half-spaces of the form

$$H_{i,j} = \left\{x \in \mathbb{R}^d\ :\ x_i \le j + 1/2\right\},\quad i = 1, \ldots, d,\ j = 1, \ldots \left\lfloor n_0^{1/d} \right\rfloor .$$

Observe that for any edge $\{x, y\} \in \binom{X}{2}$, the number of sets in $\mathcal{S}$ that crosses $\{x, y\}$ is precisely the ℓ_1-distance of x and y, which is defined as

$$\ell_1(x, y) = \sum_{i=1}^{d} |x_i - y_i|\ .$$

Using this observation, it is easy to see that for any $k \in \mathbb{N}$ and $x \in X$, the number of edges $\{x, y\}$ that is crossed by at most k sets from $\mathcal{S}$ is $O(k^d)$. Thus, there is an absolute constant c such that the total number of edges in $\binom{X}{2}$ crossed by at most k sets from $\mathcal{S}$ is at most $c \cdot nk^d$. We refer to these edges as k-*good* and denote their set with $\mathcal{G}_k$. Let $p(n) = o(n^{\alpha-1})$ and E be a uniform random sample of edges, where each $e \in \binom{X}{2}$ is picked with probability $p(n)$. Setting $k_p(n) = \left(\frac{1}{16c \cdot p(n)}\right)^{1/d}$, the expected number of $k_p(n)$-good edges in E is

$$\mathbb{E}\left[\ |E \cap \mathcal{G}_{k_p(n)}|\ \right] \leq cn\left(k_p(n)\right)^d \cdot p(n) = \frac{n}{16}\ .$$

Thus, by Markov's inequality, we have $|E \cap \mathcal{G}_{k_p(n)}| \leq \frac{n}{8}$ with probability at least $1/2$. Assume that $|E \cap \mathcal{G}_{k_p(n)}| \leq \frac{n}{8}$ holds and let $M \subset E$ be any subset of size $n/4$. Then M contains at least $n/8$ edges which are not $k_p(n)$-good. Therefore, the number of crossings between the edges of M and the sets of $\mathcal{S}$ is at least

$$\frac{n}{8} \cdot \left(\frac{1}{16c \cdot p(n)}\right)^{1/d}\ .$$

Recall that $|\mathcal{S}| = d \cdot \left\lfloor n_0^{1/d} \right\rfloor \leq dn^{1/d}$ and so by the pigeonhole principle, we get that there is a set in $\mathcal{S}$ that crosses at least

$$\frac{\frac{n}{8} \cdot \left(\frac{1}{16cp(n)}\right)^{1/d}}{|\mathcal{S}|} \geq \frac{\frac{n}{8} \cdot \left(\frac{1}{16cp(n)}\right)^{1/d}}{dn^{1/d}} = \frac{n^{1-1/d}}{8d \cdot (16c)^{1/d}} \cdot \underbrace{\left(\frac{1}{p(n)}\right)^{1/d}}_{\omega\left(n^{(1-\alpha)/d}\right)} = \omega\left(n^{1-\alpha/d}\right)$$

edges of M. This concludes the proof of Theorem 2. ◀

4 Proof of Corollary 4

Proof. It is sufficient to show that one recursive call has running time $\tilde{O}\left(n^{2+\alpha} \cdot |\mathcal{S}|\right)$ and that the picked edges $\{e_1, \ldots, e_{n/16}\}$ have expected crossing number $O\left(n^{1-\alpha/d} + \ln|\mathcal{S}|\right)$. Then the overall time complexity and crossing number guarantee follows directly by summing up these bounds over the $\log n$ recursive calls, with decreasing input sizes.
First, we bound the running time. At each iteration, we can find the lightest edge within E in time $|E| \cdot |\mathcal{S}|$, giving an overall time complexity of $O\left(n \cdot |E| \cdot |\mathcal{S}|\right) = \tilde{O}\left(n^{2+\alpha} \cdot |\mathcal{S}|\right)$.
Now we bound the expected crossing number guarantee. The key element of the proof is to bound the total weight of sets that cross the lightest edge. We introduce a function η to denote this quantity. Let $(X, \mathcal{S})$ be a set system and $\omega : \mathcal{S} \to \mathbb{R}$ be any function. Let $E \subseteq \binom{X}{2}$ be a set of edges and let $\eta(E, \omega)$ denote the length of the lightest edge in E with respect to ω, that is,

$$\eta(E, \omega) = \min_{e \in E} \sum_{S \text{ crosses } e} \omega(S).$$

We use the following observation which states that once the existence of a partial low-crossing matching is provided, we can bound $\eta(\cdot, \cdot)$ using a simple pigeonholing argument.

▶ **Observation 10.** *Let $(X, \mathcal{S})$ be a set system, $w : \mathcal{S} \to \mathbb{R}_{\geq 0}$, and κ be such that X has a (partial) matching M with crossing number at most κ with respect to $\mathcal{S}$. Then*

$$\eta(M, w) \leq \frac{w(\mathcal{S}) \cdot \kappa}{|M|}.$$

Let $E \subseteq \binom{X}{2}$ be an edge-set such that it contains a matching of size $n/4$ with crossing number $c \cdot \left(n^{1-\alpha/d} + \ln|\mathcal{S}|\right)$ for some fixed $\alpha \in (0,1]$. During the execution of RECURSIVEMWU$\left((X,\mathcal{S}),\alpha\right)$, the lightest edge at iteration i is only chosen from set $E \cap \binom{X_i}{2}$. By Theorem 1, E contains a matching M of size $|X|/4$ with crossing number $O\left(n^{1-\alpha/d}\right)$. Notice that once we remove the endpoints of e_i from X_i in iteration i, the size $|M \cap \binom{X_i}{2}|$ decreases by at most two. Thus, by the update rule (by an argument analogous to the one giving Equation (2)) and Observation 10, after $n/16$ steps, the picked edges $\{e_1 \ldots, e_{n/16}\}$ have crossing number at most

$$\begin{aligned}
\kappa_{n/16} &\leq \frac{1}{\ln 2}\left(\ln|\mathcal{S}| + \sum_{j=1}^{n/16} \frac{\eta\left(E \cap \binom{X_j}{2}, \omega_j\right)}{\omega_j(\mathcal{S})}\right) \leq \frac{1}{\ln 2}\left(\ln|\mathcal{S}| + \sum_{j=1}^{n/16} \frac{\eta\left(M \cap \binom{X_j}{2}, \omega_j\right)}{\omega_j(\mathcal{S})}\right) \\
&\leq \frac{1}{\ln 2}\left(\ln|\mathcal{S}| + \sum_{j=1}^{n/16} \frac{\omega_j(\mathcal{S}) \cdot c\left(n^{1-\alpha/d} + \ln|\mathcal{S}|\right)}{\omega_j(\mathcal{S}) \cdot \left|M \cap \binom{X_i}{2}\right|}\right) \\
&\leq \frac{1}{\ln 2}\left(\ln|\mathcal{S}| + c \cdot \sum_{j=1}^{n/16} \frac{n^{1-\alpha/d} + \ln|\mathcal{S}|}{n/4 - 2j + 2}\right) \\
&< \frac{1}{\ln 2}\left(\ln|\mathcal{S}| + c \cdot \sum_{j=1}^{n/16} \frac{n^{1-\alpha/d} + \ln|\mathcal{S}|}{n/8}\right) = O\left(n^{1-\alpha/d} + \ln|\mathcal{S}|\right).
\end{aligned}$$

◀

Remark 1: Note that Observation 10 and the fact that

> if $(X,\mathcal{S})$ has a spanning path/tree with crossing number κ with respect to $\mathcal{S}$, then any $Y \subseteq X$ has a spanning path/tree with crossing number at most 2κ with respect to $\mathcal{S}|_Y$

together imply that the recursive MWU algorithm can be used to obtain a $\log n$-approximation of spanning trees/paths of low crossing number, giving a new proof of the result of [10].

Remark 2: In order to obtain a $\log n$-approximation of optimal matchings, one needs to have a better understanding on the monotonicity of crossing numbers of optimal matchings. Even the simplest question appears to be open: if X has a perfect matching with crossing number κ with respect to $\mathcal{S}$, does any $Y \subseteq X$ has a perfect matching with crossing number at most $c \cdot \kappa$ with respect to $\mathcal{S}|_Y$? For an overview of results on the monotonicity of crossing numbers of spanning path and trees, see [18, 19].

References

1. P. K. Agarwal and J. Matoušek. On range searching with semialgebraic sets. *Discrete & Computational Geometry*, 11(4):393–418, 1994.
2. P. K. Agarwal, J. Matoušek, and M. Sharir. On range searching with semialgebraic sets. II. *SIAM Journal on Computing*, 42(6):2039–2062, 2013.
3. T. M. Chan. Optimal partition trees. *Discrete Comput. Geom.*, 47(4):661–690, 2012.
4. B. Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, New York, NY, USA, 2000.
5. B. Chazelle and E. Welzl. Quasi-optimal range searching in spaces of finite VC-dimension. *Discrete Comput. Geom.*, page 467–489, 1989.

6 M. Csikós and N. H. Mustafa. Escaping the Curse of Spatial Partitioning: Matchings with Low Crossing Numbers and Their Applications. In *37th International Symposium on Computational Geometry (SoCG 2021)*, volume 189, pages 28:1–28:17, 2021.

7 G. Ducoffe, M. Habib, and L. Viennot. Diameter computation on h-minor free graphs and graphs of bounded (distance) VC-dimension. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1905–1922, 2020.

8 S. P. Fekete, M. E. Lübbecke, and H. Meijer. Minimizing the stabbing number of matchings, trees, and triangulations. In *Proceedings of Symposium on Discrete Algorithms (SODA)*, 2004.

9 P. Giannopoulos, M. Konzack, and W Mulzer. Low-crossing spanning trees: an alternative proof and experiments. In *Proceedings of EuroCG*, 2014.

10 S. Har-Peled. Approximating spanning trees with low crossing number. *arXiv*, abs/0907.1131, 2009.

11 D. Haussler. Sphere packing numbers for subsets of the boolean n-cube with bounded Vapnik-Chervonenkis dimension. *Journal of Combinatorial Theory, Series A*, 69(2):217–232, 1995.

12 D. Haussler and E. Welzl. ϵ-nets and simplex range queries. *Discrete & Computational Geometry*, 2(2):127–151, 1987.

13 J. Komlós, J. Pach, and G. Woeginger. Almost tight bounds for ε-nets. *Discrete & Computational Geometry*, 7:163–173, 1992.

14 Y. Li, P. M. Long, and A. Srinivasan. Improved bounds on the sample complexity of learning. *J. Comput. Syst. Sci.*, 62(3):516–527, 2001.

15 J. Matoušek. Efficient partition trees. *Discrete & Computational Geometry*, 8(3):315–334, 1992.

16 J. Matoušek. *Lectures on discrete geometry*, volume 212. Springer Science & Business Media, 2013.

17 J. Matoušek, E. Welzl, and L. Wernisch. Discrepancy and approximations for bounded VC-dimension. *Combinatorica*, 13(4):455–466, 1993.

18 W. Mulzer and J. Obenaus. The tree stabbing number is not monotone. *CoRR*, abs/2002.08198, 2020. URL: `https://arxiv.org/abs/2002.08198`, `arXiv:2002.08198`.

19 J. Obenaus. *Spanning trees with low (shallow) stabbing number*. PhD thesis, ETH Zurich, 2019.

20 P. Turán. On an external problem in graph theory. *Mat. Fiz. Lapok*, 48:436–452, 1941.

21 V. Vapnik and A. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

22 E. Welzl. Partition trees for triangle counting and other range searching problems. In *Proceedings of Annual Symposium on Computational Geometry (SoCG)*, page 23–33, 1988.

23 E. Welzl. On spanning trees with low crossing numbers. In *Data Structures and Efficient Algorithms, Final Report on the DFG Special Joint Initiative*, page 233–249, 1992.