# Markov Incremental Constructions *

Bernard Chazelle
Department of Computer Science
Princeton University
35 Olden Street
Princeton, NJ 08540, USA
chazelle@cs.princeton.edu

Wolfgang Mulzer
Department of Computer Science
Princeton University
35 Olden Street
Princeton, NJ 08540, USA
wmulzer@cs.princeton.edu

## ABSTRACT

A classic result asserts that many geometric structures can be constructed optimally by successively inserting their constituent parts in random order. These randomized incremental constructions (RICs) still work with imperfect randomness: the dynamic operations need only be "locally" random. Much attention has been given recently to inputs generated by Markov sources. These are particularly interesting to study in the framework of RICs, because Markov chains provide highly nonlocal randomness, which incapacitates virtually all known RIC technology.

We generalize Mulmuley's theory of $\Theta$-series and prove that Markov incremental constructions with bounded spectral gap are optimal within polylog factors for trapezoidal maps, segment intersections, and convex hulls in any fixed dimension. The main contribution of this work is threefold: (i) extending the theory of abstract configuration spaces to the Markov setting; (ii) proving Clarkson-Shor type bounds for this new model; (iii) applying the results to classical geometric problems. We hope that this work will pioneer a new approach to average-case analysis in computational geometry.

## General Terms

Algorithms, Theory

## Keywords

Randomized incremental constructions, Expander graphs, Clarkson-Shor bound

## Categories and Subject Descriptors

F.2.2 [**Nonnumerical Algorithms and Problems**]: Geometrical problems and computations; G.3 [**Probability and Statistics**]: Markov processes

## 1. INTRODUCTION

Randomized incremental constructions (RICs) are popular for three reasons: they are widely applicable; they are as simple as one could hope; they are often optimal under random input sequences [3–6,9,10,12–15,17,20,29,30,32–34, 38–41,45]. But what if the sequences are not truly random? In the worst case, the running time typically goes up by a factor of $n$. Less obvious is the fact that perfect randomness is not actually necessary. Mulmuley proved that $O(1)$-wise independence was in fact sufficient [35]. On the other hand, Amenta et al. showed that the entropy may slowly decay during the RIC without penalty [1]; in other words, the insertion sequence can afford to be less and less random as the construction progresses. Devillers and Guigue introduced the *shuffling buffer*, which randomly permutes contiguous subsequences of the input sequence of a certain length $k$, and they provide trade-offs between the length $k$ and the running time of the RIC [16].

What these results demonstrate is that standard RIC analysis still works as long as there is sufficient *local* randomness *early* enough. Unfortunately these two features are precisely what is lacking in *Markov sources*.

What are those? A Markov source is a probabilistic model of input data that serializes the production of data over time by means of a random walk in a graph. It is widely used in queuing theory, speech recognition, gesture modeling, protein homology, computer graphics, robotics, web searching, etc. It captures the statistical correlations created by *time coherence*. In speech, for example, the randomness of the next utterance is heavily dependent on the previous ones; hence the use of hidden Markov models. In geometric applications, Markov sources have been used in ray tracing [23,43], computer games [28], robotics [19], terrain generation [44], etc. In computer science, one of the main motivations has been *locality of reference*; in particular, there exists a vast body of research in online algorithms for Markov sources [8, 22, 24–27, 36, 37, 42]. The work of Amenta et al [1] on RICs is also motivated by the desire for local access. The focus of much of modern computing has shifted over to the "data" side (partly because of the need to cope with massive data sets), and it is natural to ask what happens to a general algorithmic paradigm (RIC) when one assumes a Markov source—arguably the most widely used probabilistic model in applied science today.

Formally, a Markov chain over a finite state space $Q$ with $n$ states is an infinite sequence of random variables $X_0, X_1, \ldots$ with the following properties: (i) $X_t \in Q$ for all $t \geq 0$; (ii) $X_0$ is drawn from a given initial probability distribution $\pi_0$ over

$Q$; and (iii) there are numbers $p_{qr}$, $q, r \in Q$ such that for any $t \geq 0$ we have $\Pr[X_{t+1} = q_{t+1} | X_0 = q_0, X_1 = q_1, \ldots, X_t = q_t] = p_{q_t q_{t+1}}$, ie, the distribution of $X_{t+1}$ depends only on $X_t$. The variable $X_t$ is called the state of the Markov chain at time $t$. The $n \times n$ matrix $P$ formed by the $p_{qr}$ is called the *transition matrix* of the Markov chain. Given a graph $G = (V, E)$ with $n$ vertices and an initial probability distribution $\pi_0$ on $V$, a *random walk* on $G$ is a sequence of vertices $v_0, v_1, \ldots$, where $v_0$ is chosen according to $\pi_0$ and for every $t \geq 0$ the vertex $v_{t+1}$ is determined by following a random edge out of $v_t$. Naturally, a random walk induces a Markov chain on state space $V$.

For the purpose of this paper, our model consists of an *event graph* $G = (V, E)$ which is connected and undirected. This means that the Markov chain it defines is irreducible and reversible but not necessarily ergodic. Each node $v$ is associated with an item $x_v$ in a universe $\mathcal{U}$. Requests are specified by following a random walk, beginning at a designated start node of $G$ and hopping from node to node, each time choosing an adjacent node $v$ uniformly at random. Upon reaching $v$, item $x_v$ is inserted into the current structure. The structure in question depends, of course, on the application. In this work we consider convex hulls, trapezoidal maps, and segment intersections. The structure is the corresponding *conflict graph*. Actually, we can use the *influence graph* [3–6,17,18] or *history graph* [34], which have the advantage of supporting queries and allowing for online (semi)dynamic algorithms. This means that we do not even need to know the graph $G$ ahead of time. Our analysis, in fact, supports all known variants of RICs.

### Our results and their significance.

The main contribution of this paper is to extend Mulmuley's theory of $\Theta$-*series* [34] to Markov chains. In the course of doing so, we prove two results of independent interest: one is a generalization of the classic Clarkson-Shor counting technique for Markov sampling; the other is a new bound on mean first passage times for Markov chains with bounded spectral grap.

With the new tools we build, we are able to bound the expected complexity of RIC for convex hulls in $d$ dimensions by $O(\gamma^{-d} n^{\lfloor d/2 \rfloor} (\log n)^{\lceil d/2 \rceil})$ for $d > 3$, and $O(n(\gamma^{-1} \log n)^d)$ for $d \leq 3$, where $\gamma$ is the spectral gap of the random walk—note that $\gamma$ is a positive constant in the case of an expander. For trapezoidal maps of non-intersecting segments and segment intersections, the complexity is respectively $O(n(\gamma^{-1} \log n)^4)$ and $O((n + m)(\gamma^{-1} \log n)^7)$, where $m$ is the number of intersections.

The extra polylogs may come as a disappointment. They should not. First, they cannot be removed entirely, as can be shown by a simple cover time argument. Indeed, it is well known that even with constant spectral gap, a random walk requires $\Omega(n \log n)$ steps to visit every vertex of a graph with $n$ vertices [7]. So, one cannot hope to match the complexity bounds of the static case. The exponents we achieve are unlikely to be tight, however, and we suspect that further progress hinges on a better understanding of short-term behavior of random walks, a topic that seems to have been addressed only recently in the Markov chain literature [2].

A person familiar with the role of expanders in derandomizing **BPP** might expect that Markov sources should provide more, not less, randomizing power than, say, bounded independence. This intuition is wrong for very interesting reasons that are important to understand. The standard analysis for RICs require *global* randomness within *local* time windows. Markov sources violate that essential feature in the worst possible way. Even post-mixing, a short walk contains virtually no global randomness. Note that all previous uses of expanders for (de)randomization rely on their randomness over *global* windows: in that regard, this paper pioneers a local approach to Markov chains that is likely to find further applications. In particular, our contribution includes new results of general interest on first passage times.

That Markov RICs come within polylog and not, say, $n^\varepsilon$ of optimal is, in and of itself, a very intriguing result. In fact, we do not know any simple argument that shows that the expected RIC complexity beats that of the worst possible deterministic insertion sequence (even assuming bounded spectral gap)! It is no exaggeration to say that Markov sources shatter the foundation of RIC's analysis at its core. We show in this paper how the framework of $\Theta$-series can be partly salvaged. It is doubtful whether backward analysis can be similarly rescued, but this is a fascinating open question.

RICs provide essentially the simplest algorithms one can hope for. The message of this paper is that a tiny amount of local entropy (as provided by Markov sources) is sufficient to bring about almost all (though not quite all) the benefits of full randomization. What our work also shows is that there is nothing obvious at all about such a statement.

## 2. $\Theta$-SERIES FOR MARKOV SOURCES

For (notational) convenience, we assume that the random walk takes place on a graph that is connected and $r$-regular, for some constant $r$. The complexity of RIC is tightly coupled to the spectral gap $\gamma$, which is the difference between the first and second largest eigenvalues of the transition matrix. We use the classical notion of *configuration spaces* [34] and adapt it to the Markov model. This is done as follows: Fix a natural number $d$, the *degree* of the configuration space. Each node $v$ of $G$ is assigned an object $x_v$ chosen from a geometric universe (eg, points, hyperplanes, segments), and to each $d$-tuple $\mathbf{u} = (u_1, \ldots, u_d)$ of distinct $u_i \in V$ we assign a (possibly empty) $S_{\mathbf{u}} \subseteq V$ disjoint from $\mathbf{u}$. We denote by $f_k$ the number of $\mathbf{u}$'s such that $|S_{\mathbf{u}}| = k$ and by $f_{\leq k}$ the prefix sum $f_0 + \cdots + f_k$. We write $f_k(n)$ and $f_{\leq k}(n)$ to refer to the maximum such values over all subsets of the universe of size $n$. The coordinates of a $d$-tuple $\mathbf{u}$ play the role of the *triggers* and the sets $S_{\mathbf{u}}$ that of the *stoppers*. Naturally, $f_k$ counts the $k$-sets of the underlying range space.

The apparent simplifications of our model do not, in fact, restrict the generality of the results in any way. Indeed, our framework can just as easily handle cases where $\mathbf{u}$ is not a sequence but a multiset, where it maps to several stopper sets, or where the degree $d$ is not unique. Given a random ordered $\mathbf{u} = (u_1, \ldots, u_d)$, perform an infinite random walk from a random node in $G$. If the walk first reaches $u_1, \ldots, u_d$ in that order before hitting any node in $S_{\mathbf{u}}$, then set $\Phi = n^d |S_{\mathbf{u}}|$; else set $\Phi = 0$. Standard $\Theta$-series theory shows that the expectation of $\Phi$ determines the expected amortized complexity of RIC [34]. We postpone the proof of this result:[1]

---

[1]For convenience, we use the Vinogradov notation $\ll$ and $\gg$ for $O(\cdot)$ and $\Omega(\cdot)$, respectively.

**Master Theorem** If $f_0(n) = O(n^\alpha)$, for some constant $\alpha > 0$, then $\mathbf{E}[\Phi] \ll \gamma^{-d} n^\alpha (\log n)^{d-\alpha}$ for $\alpha > 1$ and $\mathbf{E}[\Phi] \ll \gamma^{-d} n(\log n)^d$ for $\alpha \leq 1$.

We apply the theorem to three problems: convex hulls (and hence Voronoi diagrams); trapezoidal maps of disjoint segments; and line segment intersections. The algorithms themselves operate in standard incremental fashion by inserting objects online with the help of the influence graph. The algorithms do not require knowledge of the Markov chain (which is why we do not use conflict graphs).

- CONVEX HULLS IN $\mathbb{R}^d$: The convex hull of $n$ points in $\mathbb{R}^d$ has $O(n^{\lfloor d/2 \rfloor})$ faces, which implies that $\alpha = \lfloor d/2 \rfloor$. The algorithm runs in time $O(\gamma^{-d} n^{\lfloor d/2 \rfloor} (\log n)^{\lceil d/2 \rceil})$ for $d > 3$, and $O(n(\gamma^{-1} \log n)^d)$ for $d \leq 3$.

- TRAPEZOIDAL MAPS: At each node, the trapezoidal map formed by a set of (nonintersecting) segments is maintained. The relevant configuration space is made of three subconfiguration spaces of respective degrees 2, 3, and 4. Hence, the time required by the algorithm is $O(n(\gamma^{-1} \log n)^4)$.

- SEGMENT INTERSECTIONS: The $m$ intersections among $n$ segments are computed in $O\big((n + m)(\gamma^{-1} \log n)^7\big)$ steps. The proof depends on an extension of the Master Theorem discussed in §3.

PROOF OF THE MASTER THEOREM. Recall that the *transition matrix* of a Markov process with $n$ states is the $n \times n$ matrix $P$ in which entry $(P)_{ij}$ is the probability of a transition from state $i$ to state $j$. The transition matrix of a random walk on a graph $G$ is just $G$'s normalized adjacency matrix. Furthermore, it is easy to see that for any initial probability distribution $\pi_0$, the distribution after $t$ step equals $\pi_0^T P^t$.

For technical reasons, we assume a lazy walk with transition matrix $P = \frac{1}{2}(I + M/r)$, where $M$ is the adjacency matrix of $G$. This is only for analytical convenience and an actual implementation could assume a random walk in the original graph $G$. For the cost of a constant-factor slowdown, the lazy walk brings with it well-known analytical benefits. For example, $P$ is positive semidefinite and the walk is ergodic. Fix a node $u_0 \in V$. Given any nonempty set $S \subset V$ and $u \in V \setminus S$, let $\Pr[u_0 \xrightarrow{u} S]$ be the probability that an infinite walk from $u_0$ reaches $u$ before any node in $S$, and let $t_0 = \lfloor c(1-\lambda)^{-1} \log n \rfloor$ be an upper bound on the mixing time, where $\lambda$ is the second largest eigenvalue of $P$ and $c$ is a large enough constant [11]. Note that $\lambda = 1 - \gamma/2$ and that $\lambda^{t_0} \leq 1/n$ for appropriate $c$, since

$$\ln \lambda = \ln\left(1 - \frac{\gamma}{2}\right) = -\sum_{j=1}^{\infty} \frac{1}{j}\left(\frac{\gamma}{2}\right)^j \leq -\frac{\gamma}{2} = -(1-\lambda).$$

We begin with a technical result of independent interest.

LEMMA 2.1. *For any fixed $u_0 \in V$, $S \subseteq V$, $v \in V \setminus S$,*

$$\Pr[u_0 \xrightarrow{u} S] \ll \sum_{0 \leq t < 3t_0} (P^t)_{u_0 u} + \frac{1}{(1-\lambda)|S|}.$$

PROOF. Let $Q$ be the matrix derived from $P$ by zeroing out any entry $P_{vw}$ with either $v$ or $w$ (or both) in $S \cup \{u\}$.

(We index matrix elements and vector coordinates by their corresponding nodes in $G$.) Being positive semidefinite, $Q$ has a (real) spectral decomposition $\sum_i \mu_i z_i z_i^T$ such that $\mu_1 \geq \cdots \geq \mu_n = 0$ and the $z_i$ constitute an orthonormal basis of eigenvectors. By the Perron-Frobenius theorem, $1 > \lambda$ and $1 > \mu_1$, and by the eigenvalue interlacing lemma [21], $\lambda \geq \mu_2$. By Cauchy-Schwarz, for any $v, w \in V \setminus S \cup \{u\}$

$$(Q^t)_{vw} = z_{1v} z_{1w} \mu_1^t + \sum_{i>1} z_{iv} z_{iw} \mu_i^t$$

$$\leq z_{1v} z_{1w} \mu_1^t + \mu_2^t \sqrt{\sum_{i>1} z_{iv}^2 \sum_{i>1} z_{iw}^2} \qquad (1)$$

$$\leq z_{1v} z_{1w} \mu_1^t + \lambda^t,$$

where the last inequality follows from the orthonormality of the $z_i$. Since $\mathbf{1}/\sqrt{n}$ is the principal unit eigenvector of $P$ for the eigenvalue 1, an analogous calculation for $P$ yields for any $v, w \in V$:

$$(P^t)_{vw} \leq \frac{1}{n} + \lambda^t. \qquad (2)$$

To bound $\Pr[u_0 \xrightarrow{u} S]$, we proceed as follows: First, we distinguish between short paths (with less than $3t_0$ steps) and long paths (with more than $3t_0$ steps). The contribution of the short paths constitutes the first summand in the bound of Lemma 2.1. To analyze the contribution of the long paths, we break down every long path from $u_0$ to $u$ into a pre-mixing part, a mixed portion, and the premixed part of the reverse path. We then assess the contribution of each piece. Let $N_u$ denote the set of nodes adjacent to $u$ via a nonloop edge.

$$\Pr[u_0 \xrightarrow{u} S] \leq \sum_{t < 3t_0} (P^t)_{u_0 u} + \frac{1}{r} \sum_{t \geq 3t_0} \sum_{v \in N_u} (Q^t)_{u_0 v}. \qquad (3)$$

We now break down the long paths. By (2), the last summand is bounded by

$$\frac{1}{r} \sum_{v \in N_u} \sum_{t \geq t_0} \sum_{a,b \in V} (P^{t_0})_{u_0 a} (Q^t)_{ab} (P^{t_0})_{bv}$$

$$\leq \left(\frac{1}{n} + \lambda^{t_0}\right)^2 \sum_{t \geq t_0} \sum_{a,b \in V} (Q^t)_{ab} \qquad (4)$$

$$\leq \frac{4}{n^2} \sum_{t \geq t_0} \sum_{a,b \in V} (Q^t)_{ab},$$

where the second inequality follows from the definition of $t_0$. Since $\|z_1\|_2 = 1$ and since at least $|S| + 1$ of its coordinates are zero (an easy consequence of being an eigenvector for $Q$), Cauchy-Schwarz yields $\|z_1\|_1^2 \leq n - |S| - 1$. By Perron-Frobenius, $z_1$ is nonnegative and so, $\sum_{a,b \in V} z_{1a} z_{1b} = \|z_1\|_1^2 \leq n - |S| - 1$. By (1,3,4),

$$\Pr[u_0 \xrightarrow{u} S] - \sum_{t < 3t_0} (P^t)_{u_0 u} \leq \frac{4}{n^2} \sum_{a,b \in V} \sum_{t \geq t_0} (z_{1a} z_{1b} \mu_1^t + \lambda^t)$$

$$\leq \frac{4}{n(1-\mu_1)} + \frac{4\lambda^{t_0}}{1-\lambda}. \qquad (5)$$

We need to bound $1 - \mu_1$. By an argument similar to one given by Broder and Karlin [7], we can bound $\mu_1$ away from 1: Since $z_1$ is nonnegative, $nz_1 - \|z_1\|_1 \mathbf{1}$ is normal to the

principal eigenvector of $P$ and, by Courant-Fischer,

$$\lambda \geq \frac{(nz_1 - \|z_1\|_1 \mathbf{1})^T P (nz_1 - \|z_1\|_1 \mathbf{1})}{\|nz_1 - \|z_1\|_1 \mathbf{1}\|_2^2}$$

$$\geq \frac{nz_1^T P z_1 - \|z_1\|_1^2}{n\|z_1\|_2^2 - \|z_1\|_1^2}$$

$$\geq \frac{n\mu_1 - \|z_1\|_1^2}{n - \|z_1\|_1^2}.$$

It follows that $n\mu_1 \leq n\lambda + (1-\lambda)\|z_1\|_1^2$. We just argued that $\|z_1\|_1^2 \leq n - |S| - 1$, and so $\mu_1 \leq 1 - (1-\lambda)(|S|+1)/n$, which, by (5), completes the proof. $\square$

To bound the expectation of $\Phi$, we need to understand a certain stochastic process, which we proceed to describe. A random *thread* refers either to a single node $w_1$ chosen uniformly at random (thread size of 1) or to a sequence $w_1, \ldots, w_k$ (thread size of $k > 1$), where $w_1$ is random and, for each $i > 0$, $w_{i+1}$ is the end node of a random walk from $w_i$ of length $t_i > 0$. The time sequence $\theta = (t_1, \ldots, t_{k-1})$ parameterizes the thread. Given $1 \leq \mu \leq d$, a random $\mu$-*thread* is a sequence of $\mu$ threads whose sizes add up to $d$: each thread is drawn independently and has its own size and time sequence. Its time sequence $\theta$ refers now to the collection of its constituent threads' time sequences. A $\mu$-thread forms a $d$-tuple $\mathbf{u}$ and is therefore associated with a stopper set $S_\mathbf{u}$. (We invalidate the cases where $\mathbf{u}$ has fewer than $d$ distinct nodes by setting $|S_\mathbf{u}| = \infty$.) Let $g_k^{(\mu)}$ be the probability that a random $\mu$-thread (with a given time sequence) produces $\mathbf{u}$ such that $|S_\mathbf{u}| = k$.

$$g_k^{(\mu)} = \text{Prob}\,[\,\mu\text{-thread}\, \hookrightarrow \mathbf{u}\, :\, |S_\mathbf{u}| = k\,]. \qquad (6)$$

LEMMA 2.2. *With respect to any valid time sequence, we have* $g_{\leq k}^{(\mu)} \ll (k/n)^\mu f_0(n/k)$, *for* $k > 0$.

PROOF. We use a Clarkson-Shor type argument [13] tailored for Markov chains. As usual, the idea is to sample a $1/k$ fraction of the elements. Then, we argue that a configuration $\mathbf{u}$ with $|S_\mathbf{u}| \leq k$ that appears in the sample will not be killed with high probability.

We may assume that $k \leq n/2d$, since for larger $k$, the right hand side of the bound becomes constant.

All $\mu$-threads in this proof share a given, fixed time sequence $\theta_1, \ldots, \theta_\mu$. Let $s \leq n$ be a constant to be determined later. For each $i = 1, \ldots, \mu$, pick $s$ random threads of type $\theta_i$, and define $R$ as the set of $\mathbf{u}$'s formed by taking all possible $s^\mu$ combinations of the resulting threads, one of each type. Given a fixed (nonrandom) $\mathbf{u} \in V^d$, let $p_\mathbf{u}$ denote the probability that $\mathbf{u}$ is chosen by a random $\mu$-thread. Since each starting node is chosen independently, $p_\mathbf{u}$ is of the form $\prod_{1 \leq i \leq \mu} \frac{p_{\mathbf{u},i}}{n}$, where $p_{\mathbf{u},i}$ is the probability of getting the $i$-th thread to match with the relevant part of $\mathbf{u}$, given that the first node of the $i$-th thread equals the corresponding node in $\mathbf{u}$. Therefore, $\mathbf{u}$ ends up in $R$ with probability at least $\prod_{1 \leq i \leq \mu} (1 - (1 - p_{\mathbf{u},i}/n)^s)$. Now, since $p_{\mathbf{u},i} s/n \leq 1$, we have

$$\left(1 - \frac{p_{\mathbf{u},i}}{n}\right)^s \leq 1 - \binom{s}{1}\frac{p_{\mathbf{u},i}}{n} + \binom{s}{2}\left(\frac{p_{\mathbf{u},i}}{n}\right)^2 \leq 1 - \frac{p_{\mathbf{u},i} s}{2n};$$

hence,

$$\text{Prob}[\,\mathbf{u} \in R\,] \gg p_\mathbf{u} s^\mu. \qquad (7)$$

Let $R^v$ be the collection of nodes appearing among the $d$-tuples of $R$. Given a fixed $\mathbf{u}$ with $|S_\mathbf{u}| \leq n/2d$, conditioned upon $\mathbf{u} \in R$, what is the probability that $R^v \cap S_\mathbf{u} = \emptyset$, ie, that configuration $\mathbf{u}$ is not killed? Being in $R$, $\mathbf{u}$ itself is a $\mu$-thread formed by picking exactly one thread per type out of the $s$ available ones in $R$. The $d$ nodes of $\mathbf{u}$ lie outside $S_\mathbf{u}$, so the only possibility for $R^v$ to intersect $S_\mathbf{u}$ is for any of the $(s-1)\mu$ other threads to pass through $S_\mathbf{u}$. Take one of them: it is a random walk $w_1 \cdots w_k$. The starting node $w_1$ is random, so its distribution forms an eigenvector for the thread's transition matrix with eigenvalue 1 (also true if $k = 1$). This means that each $w_i$ lies in $S_\mathbf{u}$ with probability $|S_\mathbf{u}|/n$. These events are not independent, so we use a union bound to argue that the thread $w_1 \cdots w_k$ remains outside $S_\mathbf{u}$ with probability at least $1 - d|S_\mathbf{u}|/n$. The $(s-1)\mu$ threads that are candidates for passing through $S_\mathbf{u}$ are independent, however, and thus refrain from doing so with probability at least $(1 - d|S_\mathbf{u}|/n)^{(s-1)\mu}$. Since $|S_\mathbf{u}| \leq n/2d$, we have

$$\ln\left(1 - \frac{d|S_\mathbf{u}|}{n}\right) = -\sum_{j=1}^{\infty} \frac{1}{j}\left(\frac{d|S_\mathbf{u}|}{n}\right)^j$$

$$\geq -\frac{d|S_\mathbf{u}|}{n}\sum_{j=0}^{\infty}\left(\frac{d|S_\mathbf{u}|}{n}\right)^j$$

$$\geq -2\frac{d|S_\mathbf{u}|}{n}.$$

It follows that

$$\text{Prob}[\,R^v \cap S_\mathbf{u} = \emptyset \,|\, \mathbf{u} \in R\,] \geq e^{-2d(s-1)\mu|S_\mathbf{u}|/n}.$$

If $r_\mathbf{u}$ denotes the probability that both $\mathbf{u} \in R$ and $S_\mathbf{u} \cap R^v = \emptyset$ then, by (7), setting $s = \frac{n}{dk}$ yields

$$r_\mathbf{u} = \text{Prob}[\,\mathbf{u} \in R\,] \times \text{Prob}[\,R^v \cap S_\mathbf{u} = \emptyset \,|\, \mathbf{u} \in R\,]$$

$$\gg \left(\frac{n}{dk}\right)^\mu p_\mathbf{u}\, e^{-2\mu|S_\mathbf{u}|/k};$$

therefore, since $\mu \leq d$,

$$\sum_{\mathbf{u}\,:\,|S_\mathbf{u}| \leq n/2d} r_\mathbf{u} \gg \sum_{\mathbf{u}\,:\,|S_\mathbf{u}| \leq n/2d}\left(\frac{n}{dk}\right)^\mu p_\mathbf{u}\, e^{-2d|S_\mathbf{u}|/k}$$

$$\gg \sum_{\mathbf{u}\,:\,|S_\mathbf{u}| \leq k}\left(\frac{n}{dk}\right)^\mu p_\mathbf{u}$$

$$\gg \left(\frac{n}{k}\right)^\mu g_{\leq k}^{(\mu)}.$$

Since $|R^v| \leq ds$, by definition, $|\{\mathbf{u} \in R : |S_\mathbf{u} \cap R^v| = 0\}| \leq f_0(ds)$; therefore,

$$\sum_{\mathbf{u}\,:\,|S_\mathbf{u}| \leq n/2d} r_\mathbf{u} \leq f_0(n/k).$$

Note that this holds uniformly over all time sequences for $\mu$. $\square$

We now proceed with the proof of the Master Theorem. The expectation of $\Phi$ is given by

$$\mathbf{E}\,[\Phi]$$

$$= \frac{n^d d!}{\binom{n}{d}}\sum_\mathbf{u}\frac{1}{n}\sum_{u_0 \in V}|S_\mathbf{u}|\prod_{i=0}^{d-1}\text{Pr}[u_i \overset{u_{i+1}}{\to} S_\mathbf{u} \cup \{u_{i+2}, \ldots, u_d\}],$$

where $\sum_\mathbf{u}$ ranges over all ordered subsets of $d$ distinct nodes: obviously, we may restrict the sum to $\{\mathbf{u} : |S_\mathbf{u}| > 0\}$.
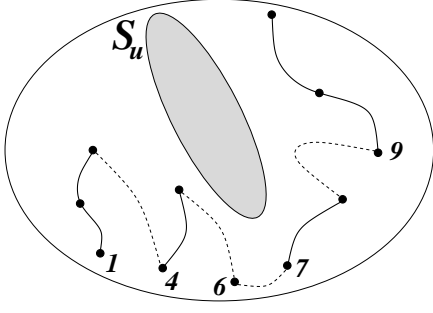
**Figure 1: The index set $L = \{1, 2, 4, 7, 9, 10\}$ defines a 5-thread.**

Note that to remove elements from $S$ cannot decrease $\Pr[u_0 \overset{u}{\to} S]$; therefore,

$$\mathbf{E}\,[\Phi] \ll \sum_{\mathbf{u}} \frac{A_{\mathbf{u}}}{n} |S_{\mathbf{u}}| \prod_{i=1}^{d-1} \Pr[u_i \overset{u_{i+1}}{\to} S_{\mathbf{u}}],$$

where, by Lemma 2.1,

$$
\begin{aligned}
A_{\mathbf{u}} &\overset{\text{def}}{=} \sum_{u_0 \in V} \Pr[u_0 \overset{u_1}{\to} S_{\mathbf{u}}] \\
&\ll \sum_{u_0 \in V} \left( \sum_{0 \le t < 3t_0} (P^t)_{u_0 u_1} + \frac{1}{(1-\lambda)|S_{\mathbf{u}}|} \right) \\
&= 3t_0 + \frac{n}{(1-\lambda)|S_{\mathbf{u}}|} \\
&\ll \frac{n}{1-\lambda} \left( \frac{1}{|S_{\mathbf{u}}|} + \frac{\log n}{n} \right).
\end{aligned}
$$

Thus, using Lemma 2.1 once more,

$$
\begin{aligned}
&\mathbf{E}\,[\Phi] \\
&\ll \frac{1}{1-\lambda} \sum_{\mathbf{u}} \left( 1 + \frac{|S_{\mathbf{u}}|\log n}{n} \right) \prod_{i=1}^{d-1} \Big\{ \sum_{t=0}^{3t_0-1} (P^t)_{u_i u_{i+1}} \quad (8) \\
&\hspace{4cm} + \frac{1}{(1-\lambda)|S_{\mathbf{u}}|} \Big\}.
\end{aligned}
$$

Writing (8) as $\mathbf{E}\,[\Phi] \ll (1-\lambda)^{-1} \sum_{\mathbf{u}} (1 + |S_{\mathbf{u}}|(\log n)/n) B_{\mathbf{u}}$, we begin with the sum $\sum_{\mathbf{u}} B_{\mathbf{u}}$. Expanding the $(d-1)$-fold product $B_{\mathbf{u}}$ produces $2^{d-1}$ terms of the form

$$\left( \frac{1}{1-\lambda} \right)^j \frac{1}{|S_{\mathbf{u}}|^j} \prod_{i \in L} \sum_{t < 3t_0} (P^t)_{u_i u_{i+1}}, \qquad (9)$$

where $L \subseteq [d-1]$ and $j + |L| = d - 1$. The index set $L$ specifies the parameters of a $\mu$-thread (except for its time sequence). Indeed, break $[d]$ into $\mu = j+1$ intervals by applying the rule that $i+1$ is in the same interval as any $i \in L$. In Figure 1, $d = 11$, $\mu = 5$, $j = 4$, $L = \{1, 2, 4, 7, 9, 10\}$, and the threads are $[1,3], [4,5], [6], [7,8], [9,10,11]$. All we can say about the time sequences is that the total number of elements $t_1, t_2, \ldots$ in all of them is exactly $|L| = d - \mu$. So, while summing (9) over all $\mathbf{u}$, let us first fix the parameters of the $\mu$-thread, with the understanding that the upper bound we derive will be off by at most a factor of $(3t_0)^{d-\mu}$. We use the superscripts $\mu, \theta$ in the sums to indicate a fixed $\mu$ or a fixed time sequence $\theta$ (or both).

$$\sum_{\mathbf{u}}^{\mu,\theta} (9) \le \left( \frac{1}{1-\lambda} \right)^j \sum_{\mathbf{u}}^{\mu,\theta} \frac{n^\mu}{|S_{\mathbf{u}}|^j} \Pr[\,\mu\text{-thread} = \mathbf{u}\,]. \quad (10)$$

Note the presence of the factor $n^\mu$ to make up for the fact that $(P^t)_{u_i u_{i+1}}$ is a conditional probability. Assume that $j > 0$. Using summation by parts, we get that the sum $\sum_{\mathbf{u}}^{\mu,\theta} |S_{\mathbf{u}}|^{-j} \Pr[\mu\text{-thread} = \mathbf{u}]$ is at most

$$
\begin{aligned}
\sum_{k=1}^{n} \frac{g_k^{(\mu)}}{k^j} &= \sum_{k=1}^{n} g_{\le k}^{(\mu)} \left( \frac{1}{k^j} - \frac{1}{(k+1)^j} \right) + \frac{g_{\le n}^{(\mu)}}{n^j} \\
&\ll \sum_{k=1}^{n} \frac{g_{\le k}^{(\mu)}}{k^{j+1}} + \frac{g_{\le n}^{(\mu)}}{n^j} \qquad (11) \\
&\ll \frac{1}{n^\mu} \sum_{k=1}^{n} \frac{f_0(n/k)}{k^{j+1-\mu}} + \frac{1}{n^j},
\end{aligned}
$$

where Lemma 2.2 is used for the last inequality. By (10) and using the identity $\mu = j + 1$,

$$
\begin{aligned}
\sum_{\mathbf{u}}^{\mu,\theta} (9) &\ll \left( \frac{1}{1-\lambda} \right)^j \left( \sum_{k=1}^{n} f_0(n/k) + \frac{1}{n^j} \right) \qquad (12) \\
&\ll \left( \frac{1}{1-\lambda} \right)^j \sum_{k=1}^{n} \left( \frac{n}{k} \right)^\alpha ;
\end{aligned}
$$

Multiplying by the corrective factor $(3t_0)^{d-\mu}$,

$$\sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} \ll \left( \frac{1}{1-\lambda} \right)^{d-1} (\log n)^{d-\mu} \sum_{k=1}^{n} \left( \frac{n}{k} \right)^\alpha. \quad (13)$$

We can now easily cover all cases:

**(I)** $j > 0$ and $\alpha \le 1$: $\sum_k (n/k)^\alpha = O(n \log n)$; hence, using $\mu = j + 1$,

$$
\begin{aligned}
\sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} &\ll (1-\lambda)^{1-d} n (\log n)^{d-j-1} \\
&\ll (1-\lambda)^{1-d} n (\log n)^{d-2}.
\end{aligned}
$$

**(II)** $j > 0$ and $\mu \ge \alpha \ge 2$: $\sum_k (n/k)^\alpha \ll n^\alpha$; hence

$$
\begin{aligned}
\sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} &\ll (1-\lambda)^{1-d} n^\alpha (\log n)^{d-\mu} \\
&\ll (1-\lambda)^{1-d} n^\alpha (\log n)^{d-\alpha}.
\end{aligned}
$$

**(III)** $\alpha > \mu$ or $j = 0$: by (10), $\sum_{\mathbf{u}}^{\mu,\theta} (9) \ll (1-\lambda)^{-j} n^\mu$; hence

$$\sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} \ll (1-\lambda)^{1-d} n^\mu (\log n)^{d-1-j}.$$

If $\alpha > \mu$, then $\sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} = o(n^\alpha)$. If $j = 0$, then $\sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} \ll (1-\lambda)^{1-d} n (\log n)^{d-1}$.

Ignoring time sequences, the number of $\mu$-thread types depends only on $d$ and is bounded by a constant. We conclude that

$$\sum_{\mathbf{u}} B_{\mathbf{u}} \ll (1-\lambda)^{1-d} \max \left\{ n(\log n)^{d-1},\ n^\alpha (\log n)^{d-\alpha} \right\}.$$

Going back to (8), recall that

$$\mathbf{E}\left[\Phi\right] \ll (1-\lambda)^{-1} \sum_{\mathbf{u}} (1 + |S_{\mathbf{u}}|(\log n)/n) B_{\mathbf{u}}.$$

To handle $\frac{\log n}{n} \sum_{\mathbf{u}} B_{\mathbf{u}} |S_{\mathbf{u}}|$, we revisit the above calculation. With the additional $\frac{\log n}{n} |S_{\mathbf{u}}|$ factor, Equation (9) now becomes

$$\frac{\log n}{n} \left(\frac{1}{1-\lambda}\right)^{j} \frac{1}{|S_{\mathbf{u}}|^{j-1}} \prod_{i \in L} \sum_{t < 3t_0} (P^t)_{u_i u_{i+1}}.$$

We fist consider the case $j \geq 1$. Retracing our steps through Equations (10,11,12), we get the analogue to (13):

$$\frac{\log n}{n} \sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} |S_{\mathbf{u}}| \ll \left(\frac{1}{1-\lambda}\right)^{d-1} (\log n)^{d+1-\mu} \sum_{k=1}^{n} \left(\frac{n}{k}\right)^{\alpha-1}.$$

If $\alpha \leq 1$, using $\mu = j + 1$, we get

$$\frac{\log n}{n} \sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} |S_{\mathbf{u}}| \ll (1-\lambda)^{1-d} (\log n)^{d-j} n$$
$$\leq (1-\lambda)^{1-d} (\log n)^{d-1} n,$$

and if $\alpha > 2$, then clearly $\frac{\log n}{n} \sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} |S_{\mathbf{u}}| = o(n^{\alpha})$. Finally, if $j = 0$, then by $\frac{\log n}{n} \sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} |S_{\mathbf{u}}| \leq \log n \sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}}$ and Case **(III)** above, it follows that $\frac{\log n}{n} \sum_{\mathbf{u}}^{\mu} B_{\mathbf{u}} |S_{\mathbf{u}}| \ll (1-\lambda)^{1-d} n(\log n)^d$.
Thus,

$$\mathbf{E}\left[\Phi\right] \ll (1-\lambda)^{-1} \max \left\{ n(\log n)^d, \, n^{\alpha} (\log n)^{d-\alpha} \right\}.$$

This completes the proof of the Master Theorem. □

# 3. EXTENSIONS

The Master Theorem cannot be used for the trapezoidal map of intersecting segments. The reason is that the complexity of an arrangement of $n$ segments depends on both $n$ and the number $m$ of intersections. We show how to extend the Master Theorem to handle this case. The problem can be described by a configuration space of degree 7 and $f_0(n,m) = O(n+m)$. We need to strengthen Lemma 2.2:

LEMMA 3.1. *Let $g_{\leq k}^{(\mu)}$ be as defined in (6); with respect to any valid time sequence, for any $k > 0$,*

$$g_{\leq k}^{(\mu)} \ll (k/n)^{\mu} O\left((n+m)/k\right).$$

PROOF. We use the same notation as in Lemma 2.2. We only need a better upper bound on $\sum_{\mathbf{u}} r_{\mathbf{u}}$, the expected complexity of the trapezoidal decomposition of the sample $R^v$. To do this, we bound the expected number of intersections among the line segments in $R^v$. Let $I$ be an intersection and let $x$ be one of its defining segments: $I$ can only be present in the trapezoidal decomposition of $R^v$ if $x \in R^v$. This happens with probability $1 - (1-d/n)^s = \Theta(1/k)$, since $ds/n \leq 1$. By linearity of expectation, it follows that the expected number of intersections in $R^v$ is $O(m/k)$ and we get the desired upper bound of $O((n+m)/k)$ on the complexity of the trapezoidal decomposition of $R^v$. Together with the lower bound from the proof of Lemma 2.2, this completes the proof. □

The desired result follows now by repeating the proof of the Master Theorem with the bound $g_{\leq k}^{(\mu)} \ll (n+m)/k$

in (11). To summarize, the $m$ intersections among $n$ segments are computed in time $O\left((n+m)(\gamma^{-1} \log n)^7\right)$, as we claimed earlier. □

*Revisiting the Clarkson-Shor bound.*

While proving the Master Theorem, we obtained a variant of the Clarkson-Shor bound suited for our Markov model (Lemma 2.2). We believe that this lemma is of independent interest and could lead to new bounds on the number of $k$-sets when certain restrictions on the defining elements are imposed. Here is a toy example: Let $P$ be a set of $n$ points in $\mathbb{R}^3$. Let $H$ be the set of planes in $\mathbb{R}^3$ spanned by triplets of the form $(x, y, \mathrm{nn}(x))$ for $x, y \in P$, where $\mathrm{nn}(x)$ denotes the nearest neighbor of $x$ in $P$. A plane $h \in H$ *conflicts* with a point $p \in P$ if $p$ lies below $h$. Let $f_{\leq k}$ denote the number of planes in $H$ that conflict with at most $k$ points.

COROLLARY 3.2. $f_{\leq k} = O(nk)$.

Compare this with the well-known Clarkson-Shor bound of $O(nk^2)$ for the unrestricted case.

PROOF. We define an event graph $G$ by choosing the nearest neighbor graph of $P$, ie, the undirected graph whose vertices are the points in $P$ and in which there is an edge between $p, q \in P$ if $p$ is the nearest neighbor of $q$ in $P$ or vice versa. It is well known that $G$ has bounded degree (eg, [31]). We choose $d = 3$ and $\mu = 2$. The first thread has size two with time sequence (1), the second thread has size one, where the first node of each thread is sampled according to the stationary distribution of $G$. In other words, the sampling is defined as follows: Let $r$ be the maximum degree of $G$, and let $m$ be the number of edges in $G$. Pick a random node in $G$, where node $v$ is chosen with probability $\deg(v)/2m$ and take one random step, and then pick another random node, again choosing node $v$ with probability $\deg(v)/2m$. This yields a triplet of points that defines a plane in $H$, and each triplet appears with probability $\Theta(1/n^2)$. By Lemma 2.2, it now follows that the probability that we sample a plane that conflicts with at most $k$ points is $O((n/k)(k/n)^2)$, as $f_0(n) = O(n)$, which proves the claim.

Technically, Lemma 2.2 applies only to regular graphs, while $G$ has bounded, but possibly varying, degree. But as we said earlier, our discussion easily generalizes to the nonregular case—at a loss of only a constant factor. Indeed, consider the proof of Lemma 2.2. We may assume that $k \leq n/2dr$. Set $s = n/dkr$. Then Equation (7) still holds, since for a given $\mathbf{u} \in V^d$, the probability $p_{\mathbf{u}}$ that $\mathbf{u}$ is chosen by a random $\mu$-thread is now of the form $\prod_{1 \leq i \leq \mu} \frac{d_i p_{\mathbf{u},i}}{2m}$, where $d_i$ is the degree of the first node of $\mathbf{u}$ that corresponds to the $i$-th thread, and by our definition of $s$ we have $d_i p_{\mathbf{u},i} s/2m \leq r p_{\mathbf{u},i} s/2n \leq 1$.

Next, we need to bound the probability that $\mathbf{u}$ with $|S_{\mathbf{u}}| \leq n/2dr$ is not killed, given that it is present in $R$. Since we sample according to the stationary distribution, each node of a $\mu$-thread lies in $S_{\mathbf{u}}$ with probability at most $r|S_{\mathbf{u}}|/n$. Proceeding as before, we now get

$$\mathrm{Prob}[R^v \cap S_{\mathbf{u}} = \emptyset \,|\, \mathbf{u} \in R] \geq e^{-2d(s-1)\mu r |S_{\mathbf{u}}|/n}.$$

and

$$r_{\mathbf{u}} \gg \left(\frac{n}{drk}\right)^{\mu} p_{\mathbf{u}} e^{-2d|S_{\mathbf{u}}|/k}.$$

Thus, as before,

$$\sum_{\mathbf{u}:\,|S_{\mathbf{u}}|\leq n/2dr} r_{\mathbf{u}} \gg \left(\frac{n}{k}\right)^{\mu} g_{\leq k}^{(\mu)}$$

and

$$\sum_{\mathbf{u}:\,|S_{\mathbf{u}}|\leq n/2dr} r_{\mathbf{u}} \leq f_0(ds) = f_0(n/rk) \leq f_0(n/k),$$

by our assumption that $f_0(n) = n^{\alpha}$. This finishes the extension of Lemma 2.2 to the bounded degree case. $\square$

## Acknowledgments

## 4. REFERENCES

[1] N. Amenta, S. Choi, and G. Rote. Incremental constructions con BRIO. In *SCG '03: Proceedings of the nineteenth annual symposium on Computational geometry*, pages 211–219, New York, NY, USA, 2003. ACM.

[2] G. Barnes and U. Feige. Short random walks on graphs. *SIAM J. Discrete Math.*, 9(1):19–28, 1996.

[3] J.-D. Boissonnat, O. Devillers, R. Schott, M. Teillaud, and M. Yvinec. Applications of random sampling to on-line algorithms in computational geometry. *Discrete Comput. Geom.*, 8(1):51–71, 1992.

[4] J. D. Boissonnat and M. Teillaud. The hierarchical representation of objects: the delaunay tree. In *SCG '86: Proceedings of the second annual symposium on Computational geometry*, pages 260–268, New York, NY, USA, 1986. ACM.

[5] J.-D. Boissonnat and M. Teillaud. On the randomized construction of the Delaunay tree. *Theoret. Comput. Sci.*, 112(2):339–354, 1993.

[6] J.-D. Boissonnat and M. Yvinec. *Algorithmic geometry*. Cambridge University Press, New York, NY, USA, 1998.

[7] A. Z. Broder and A. R. Karlin. Bounds on the cover time. *J. Theoret. Probab.*, 2(1):101–120, 1989.

[8] P. Chassaing. Optimality of move-to-front for self-organizing data structures with locality of references. *Ann. Appl. Probab.*, 3(4):1219–1240, 1993.

[9] O. Cheong, K. Mulmuley, and E. A. Ramos. Randomization and derandomization. In J. E. Goodman and J. O'Rourke, editors, *Handbook of discrete and computational geometry*, chapter 40, pages 895–926. CRC Press, Inc., Boca Raton, FL, USA, 2nd edition, 2004.

[10] L. P. Chew. Building voronoi diagrams for convex polygons in linear expected time. Technical Report PCS-TR90-147, Dartmouth College, Hanover, NH, USA, 1990.

[11] F. R. K. Chung. *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92)*. American Mathematical Society, Providence, RI, USA, 1997.

[12] K. L. Clarkson, K. Mehlhorn, and R. Seidel. Four results on randomized incremental constructions. *Comput. Geom.*, 3(4):185–212, 1993.

[13] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry. II. *Discrete Comput. Geom.*, 4(5):387–421, 1989.

[14] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications*. Springer Verlag, 2000.

[15] O. Devillers. The Delaunay hierarchy. *Internat. J. Found. Comput. Sci.*, 13:163–180, 2002. special issue on triangulations.

[16] O. Devillers and P. Guigue. The shuffling buffer. *Int. J. Comput. Geometry Appl.*, 11(5):555–572, 2001.

[17] O. Devillers, S. Meiser, and M. Teillaud. Fully dynamic Delaunay triangulation in logarithmic expected time per operation. *Comput. Geom.*, 2(2):55–80, 1992.

[18] K. Dobrindt and M. Yvinec. Remembering conflicts in history yields dynamic algorithms. In *Algorithms and computation (Hong Kong, 1993)*, volume 762 of *Lecture Notes in Comput. Sci.*, pages 21–30, Berlin, 1993. Springer.

[19] D. Fox, W. Burgard, and S. Thrun. Markov localization for reliable robot navigation and people detection. In *Selected Papers from the International Workshop on Sensor Based Intelligent Robots*, pages 1–20, London, UK, 1999. Springer-Verlag.

[20] L. J. Guibas, D. E. Knuth, and M. Sharir. Randomized incremental construction of Delaunay and Voronoi diagrams. *Algorithmica*, 7(4):381–413, 1992.

[21] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990.

[22] G. Hotz. Search trees and search graphs for markov sources. *Elektronische Informationsverarbeitung und Kybernetik*, 29(5):283–292, 1993.

[23] H. W. Jensen. *Realistic image synthesis using photon mapping*. A K Peters Ltd., Natick, MA, 2001.

[24] S. Kapoor and E. M. Reingold. Stochastic rearrangement rules for self-organizing data structures. *Algorithmica*, 6(2):278–291, 1991.

[25] A. R. Karlin, S. J. Phillips, and P. Raghavan. Markov paging. *SIAM J. Comput.*, 30(3):906–922, 2000.

[26] L. K. Konneker and Y. L. Varol. A note on heuristics for dynamic organization of data structures. *Inf. Process. Lett.*, 12(5):213–216, 1981.

[27] K. Lam, M. Y. Leung, and M. K. Siu. Self-organizing files with dependent accesses. *J. Appl. Probab.*, 21(2):343–359, 1984.

[28] P. Lu, X. Zeng, X. Huang, and Y. Wang. Navigation in 3D game by markov model based head pose estimating. In *ICIG '04: Proceedings of the Third International Conference on Image and Graphics*, pages 493–496, Washington, DC, USA, 2004. IEEE Computer Society.

[29] J. Matoušek, M. Sharir, and E. Welzl. A subexponential bound for linear programming. *Algorithmica*, 16(4-5):498–516, 1996.

[30] K. Mehlhorn, M. Sharir, and E. Welzl. Tail estimates for the efficiency of randomized incremental algorithms for line segment intersection. *Comput. Geom.*, 3:235–246, 1993.

[31] G. L. Miller, S.-H. Teng, W. Thurston, and S. A. Vavasis. Separators for sphere-packings and nearest

neighbor graphs. *J. ACM*, 44(1):1–29, 1997.

[32] K. Mulmuley. A fast planar partition algorithm. I. *J. Symbolic Comput.*, 10(3-4):253–280, 1990.

[33] K. Mulmuley. A fast planar partition algorithm. II. *J. ACM*, 38(1):74–103, 1991.

[34] K. Mulmuley. *Computational Geometry: An Introduction through Randomized Algorithms.* Prentice-Hall, Englewood Cliffs, 1994.

[35] K. Mulmuley. Randomized geometric algorithms and pseudorandom generators. *Algorithmica*, 16(4-5):450–463, 1996.

[36] R. M. Phatarfod, A. J. Pryde, and D. Dyte. On the move-to-front scheme with Markov dependent requests. *J. Appl. Probab.*, 34(3):790–794, 1997.

[37] F. Schulz and E. Schömer. Self-organizing data structures with dependent accesses. In *ICALP*, pages 526–537, 1996.

[38] O. Schwarzkopf. Dynamic maintenance of geometric structures made easy. In *Proceedings of the 32nd annual symposium on Foundations of computer science*, pages 197–206, Los Alamitos, CA, USA, 1991. IEEE Computer Society Press.

[39] R. Seidel. A simple and fast incremental randomized algorithm for computing trapezoidal decompositions and for triangulating polygons. *Comput. Geom.*, 1:51–64, 1991.

[40] R. Seidel. Small-dimensional linear programming and convex hulls made easy. *Discrete Comput. Geom.*, 6(5):423–434, 1991.

[41] R. Seidel. Backwards analysis of randomized geometric algorithms. In *New trends in discrete and computational geometry*, volume 10 of *Algorithms Combin.*, pages 37–67. Springer, Berlin, 1993.

[42] G. S. Shedler and C. Tung. Locality in page reference strings. *SIAM Journal on Computing*, 1(3):218–241, 1972.

[43] E. Veach and L. J. Guibas. Metropolis light transport. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 65–76, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[44] C. Wellington, A. Courville, and A. T. Stentz. A generative model of terrain for autonomous navigation in vegetation. *Int. J. Rob. Res.*, 25(12):1287–1304, 2006.

[45] E. Welzl. Smallest enclosing disks (balls and ellipsoids). In *New results and new trends in computer science (Graz, 1991)*, volume 555 of *Lecture Notes in Comput. Sci.*, pages 359–370. Springer, Berlin, 1991.