One line and n points*

Bernd Gärtner† Falk Tschirschnitz[†] Emo Welzl
Institut für Theoretische Informatik, ETH Zürich
ETH Zentrum, CH-8092 Zürich, Switzerland
{gaertner,tschirsc,welzl}@inf.ethz.ch

József Solymosi[‡]
Department of Mathematics, University of British Columbia
1984 Mathematics Road, Vancouver, BC, Canada V6T 1Z2
solymosi@math.ubc.ca

Pavel Valtr
Department of Applied Mathematics, Charles University, Prague
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
valtr@kam.mff.cuni.cz

Abstract

We analyze a randomized pivoting process involving one line and n points in the plane. The process models the behavior of the Random-Edge simplex algorithm on simple polytopes with n facets in dimension $n-2$. We obtain a tight $O(\log^2 n)$ bound for the expected number of pivot steps. This is the first nontrivial bound for Random-Edge which goes beyond bounds for specific polytopes. The process itself can be interpreted as a simple algorithm for certain 2-variable linear programming problems, and we prove a tight $\Theta(n)$ bound for its expected runtime.

*A preliminary version of this paper has appeared in *Proc. 33rd Ann. ACM Symp. on the Theory of Computing (STOC)* (2001) 306-315.

†Part of this work was done while the author was supported by the Swiss Science Foundation (SNF), project No. 21-50647.97.

‡Supported by the joint Berlin/Zürich graduate program Combinatorics, Geometry, Computation, financed by German Science Foundation (DFG) and ETH Zürich. On leave from MTA SZTAKI.

Keywords: Simplex algorithm, linear programming, Random-edge pivot rule, randomization.

1 Introduction

Let S be a set of n points in general position in the plane (i.e., no three on a common line), and let ℓ be a vertical line which is disjoint from S and from all intersections of segments connecting pairs of points in S. We use $e \in \binom{S}{2}$ for a pair of points and $\overline{e}$ for the segment $\operatorname{conv} e$.[1] e is called an ℓ-edge if ℓ separates its endpoints.

For a non-vertical line λ, $\operatorname{below}(\lambda)$ denotes the set of points from S below λ. For an edge e, we use $\operatorname{below}(e)$ for the set of points in S that lie below the line carrying $\overline{e}$. In fact, we will frequently use 'below $\overline{e}$' short for 'below the line carrying $\overline{e}$'. Given an ℓ-edge e and $p \in \operatorname{below}(e)$, we use $\operatorname{pivot}(e,p)$ for the unique ℓ-edge $\{p,q\}$, $q \in e$ (see Figure 1(a)).

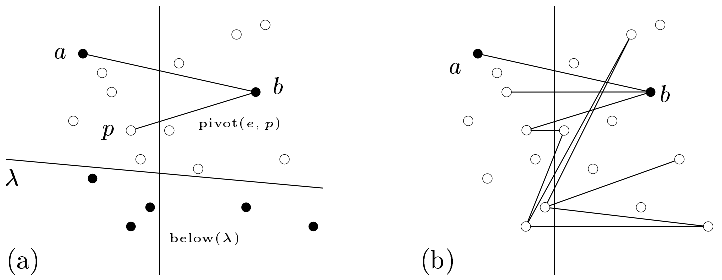


Figure 1: (a) The setup and (b) a pivoting sequence.

We are interested in the randomized process given in Figure 2 which we will call *fast process*. It finds the unique ℓ-edge e such that $\operatorname{below}(e) = \emptyset$, given some initial ℓ-edge $\{a,b\}$.[2]

Figure 1(b) shows a possible sequence of edges that may arise during the process.

The significance of this process becomes apparent when we interpret it in terms of the simplex method for linear programming. It models the

[1] Given a set of points S, $\operatorname{conv} S$ denotes the *convex hull* of S.
[2] "$p \leftarrow_{\text{random}} S$;" means *$p$ is chosen uniformly at random from the set S.*

$e \leftarrow \{a, b\}$;
while $\text{below}(e) \neq \emptyset$ **do**
  $p \leftarrow_{\text{random}} \text{below}(e)$;
  $e \leftarrow \text{pivot}(e, p)$;

Figure 2: Fast process.

behavior of the RANDOM-EDGE simplex algorithm on any linear program whose feasible region is a simple n-facet polytope in dimension $n-2$. Here, a pivot step corresponds to a move from one vertex to an adjacent vertex along an incident edge chosen at random among the ones that improve the objective function. (Cf. the appendix for a detailed explanation.)

Despite the fact that RANDOM-EDGE is the most natural randomized variant of the simplex method, and that it has received considerable attention ([11], [16, Exercise 8.10*], [15, Section 9.10], [1], [5], [13]), nontrivial bounds for its expected performance could only be derived for specific instances [5]. In particular, the case of general n-facet polytopes in dimension d could only be handled for $n = d+1$ (when the polytope is a simplex—we come back to this case at the beginning of the next section). If $n = d+2$, as in our scenario, only the trivial $O(n^2)$ bound—obtained from the maximum number of vertices a polytope with these parameters can have—was known.

In this paper, we obtain two new results. First, we prove a tight bound of $O(\log^2 n)$ for the number of pivot steps needed by the randomized process of Figure 2 (equivalently, by RANDOM-EDGE on an n-facet polytope in dimension $n-2$). This is an exponential improvement over the previous bound.

Second, we analyze a variant of this process which we call the *slow process*. Here, each round chooses a random point among *all* points and only performs a pivot step if the chosen point is below the current edge (Figure 3).

From the analysis of the fast process, we get an $O(n \log^2 n)$ bound for the expected number of iterations of the slow process. We can prove that the bound is actually $\Theta(n)$, making the slow process an ultimately simple and asymptotically optimal algorithm for 2-variable LP that can be transformed to the following linear program (LP) in two variables k, d:

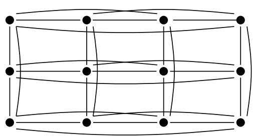
$$\begin{array}{ll} \text{maximize} & d \\ \text{subject to} & y_i \geq kx_i + d, \quad i = 1 \ldots n. \end{array} \tag{1}$$

$e \leftarrow \{a, b\}$;
while below(e) $\neq \emptyset$ **do**
  $p \leftarrow_{\text{random}} S$;
  if p below $\overline{e}$ **then**
    $e \leftarrow$ pivot(e, p);

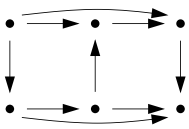
Figure 3: Slow process.

If ℓ is the y-axis, this is just the problem of finding the lowest ℓ-edge determined by points $p_i = (x_i, y_i), i = 1, \ldots, n$. (Every feasible solution of the LP (1) corresponds to a line with all points above. The corresponding value d is the height of the intersection point of this line with ℓ. It follows that the optimal line supports the unique edge of the lower convex hull intersecting ℓ which, therefore, must be an ℓ-edge.) An alternative way to look at the fast process is due to Dantzig ([3]): he used it as the geometric interpretation of the *dual* of the LP (1), calling it the *column geometry*.

Finally, we want to mention that one can interpret both processes as random walks on the directed graph whose nodes are the ℓ-edges, and whose arcs are the pairs $(e, \text{pivot}(e, p))$, for all ℓ-edges e and points p below e. The underlying undirected graph is what we call a *grid graph* with n_L rows and n_R columns, where n_L (n_R, resp.) is the number of points to the left (right, resp.) of ℓ (see Figure 4(a)).



(a)



(b)

Figure 4: (a) A grid graph with 3 rows and 4 columns and (b) the forbidden subgraph.

After pairing rows (columns) with points on the left (right) of ℓ in an arbitrary fashion, every node corresponds to some ℓ-edge, and the graph receives its orientation. Gärtner et.al. observed ([6], [7]) that if such a grid orientation is induced by a line and n points, then it is acyclic, every subgrid[3] has a unique sink, and no subgrid is isomorphic to the graph shown in Figure 4(b).

On the other hand, a grid orientation is called *admissible* if it has these properties. As shown in [7], any admissible grid orientation is induced by a line ℓ and n points in the sense that we find *pseudolines* through the pairs of points whose relative orders along ℓ determine the orientation, see Figure 5. Note that the arguments used in the proofs below are also valid in this general setting. In fact, one can translate the proofs into pure combinatorial arguments on admissible grid orientation ([17]).

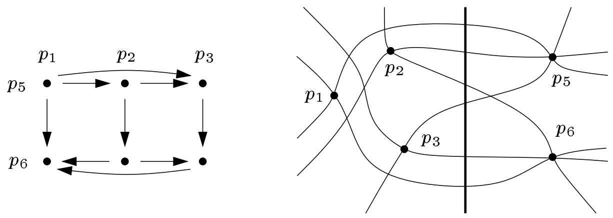


Figure 5: A grid orientation and its realization.

2 Upper Bounds

Here we analyze the expected number of pivots in the process described in Figure 2.

Warm-up. It is instructive to first have a look at the 1-dimensional counterpart as given in Figure 6. For that we are given a set M of real numbers, and $a \in M$. As above, this process can be interpreted in terms of the simplex algorithm; this time, it models the behavior of the RANDOM-EDGE variant on n-facet polytopes in dimension $n-1$, hence on simplices.

It is not difficult[4] to show that the expected number of while-loops performed is exactly the Harmonic number H_{m-1}, where m is the rank,

[3]A graph induced by all the nodes in some of the rows and columns.

[4]For $m \in \mathbf{N}$, let f_m denote the expected number of iterations when starting with an element of rank m. Then $f_1 = 0$ and $f_{m+1} = 1 + \frac{1}{m}\sum_{i=1}^{m} f_i$. Now $f_m = H_{m-1}$ follows.

$r \leftarrow a$;
while $r \neq \min(M)$ **do**
 $r \leftarrow_{\text{random}} \{x \in M \mid x < r\}$;

Figure 6: 1-dimensional fast process.

$1+|\{x \in M \mid x < a\}|$, of a in M. But we want to provide a rough estimate in the spirit of the later analysis instead. Let X_i, $i \in \mathbf{N}_0$, be the random variable for the number of iterations of the while-loop with

$$2^i \leq |\{x \in M \mid x < \rho\}| < 2^{i+1} \tag{2}$$

for ρ the value of r at the beginning of the respective iteration of the while-loop. For $m \geq 2$, the random variable $Z = \sum_{i=0}^{\lfloor \log_2(m-1) \rfloor} X_i$ gives the overall number of executions of the while-loop. $\mathrm{E}(X_i) \leq 2$ for all $i \in \mathbf{N}_0$, since, whenever (2) holds, we have a chance of at least $\frac{1}{2}$ to choose an element of rank 2^i or smaller.
Hence, $\mathrm{E}(Z) \leq 2(1 + \lfloor \log_2(m-1) \rfloor) = O(\log m)$.

The obvious extension of that analysis to the 2-dimensional process fails, since the number of points below edges appearing in the process oscillates. In fact, the number of points below the current edge is no measure of progress at all. This number may be 1, we pivot, and the number becomes as large as $n-3$ (see Figure 7).

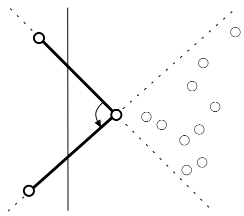


Figure 7: We thought we were so close!

k-Lines as Milestones. Here is the crucial definition that will allow us to measure progress. Given $k \in \mathbf{N}_0$, a non-vertical line λ is called *a k-line of S and ℓ* if on both sides of ℓ there are exactly k points from S below λ.

It is easy to see that every point $x \in \ell$ is contained in a k-line for some $k \in \mathbf{N}_0$, as long as x is disjoint from all segments connecting two points in S. Start with a line through x that has large slope so that all points on the right side of ℓ are below, and all on the left side are above. Now rotate the line by decreasing its slope. Eventually, we will reach the situation opposite to what we started with: No points below to the right, all below to the left. All transitions in between change the number of points below on exactly one side by ± 1. Somewhere in between we must have had a transition where the numbers of points below were the same on both sides.

A k-line disjoint from S exists for all k,

$$0 \le k \le m := \min\{n_{\mathrm{L}}, n_{\mathrm{R}}\},$$

n_{L} the number of points in S left of ℓ, and $n_{\mathrm{R}} := n - n_{\mathrm{L}}$. For each i, $0 \le i \le \lfloor \log_2 m \rfloor$, fix some 2^i-line λ_i disjoint from S. Line λ_0 has to be chosen, so that the only edge intersecting ℓ below λ_0 is the edge ε with $\mathrm{below}(\varepsilon) = \emptyset$. Moreover, let $\lambda_{\lfloor \log_2 m \rfloor + 1}$ be some m-line that intersects ℓ above all ℓ-edges (and above[5] $\lambda_{\lfloor \log_2 m \rfloor}$). The line λ_i intersects ℓ below λ_j for $0 \le i < j \le \lfloor \log_2 m \rfloor + 1$.

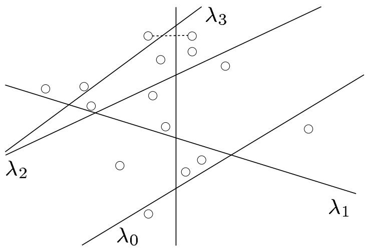


Figure 8: Setting milestones.

We define the random variable X_i, $i = 0, 1, \ldots, \lfloor \log_2 m \rfloor$, as the number of executions of the while-loop (in Figure 2) where the current segment $\overline{\varepsilon}$ intersects ℓ below λ_{i+1} but not below λ_i (ε is the value of e at the beginning of the respective iteration of the while-loop). The sequence of these executions we call *phase i of the process*[6]. The careful choice of λ_0 ensures that completion of phase 0 entails completion of the whole process. Hence, $Z = \sum_{i=0}^{\lfloor \log_2 m \rfloor} X_i$ is the random variable whose expectation we want to analyze.

[5] This is automatically satisfied, unless m is a power of 2.

[6] Note that phases count *down* during the process.

We will show that $\mathrm{E}(X_i) = O(\log n)$ for all i and, hence,
$\mathrm{E}(Z) = O((\log n)(1 + \log m)) = O(\log^2 n)$.

Analysis of a Single Phase. Fix some i, $0 \le i \le \lfloor \log_2 m \rfloor$, set $k = 2^i$, $\lambda' = \lambda_i$ and $\lambda = \lambda_{i+1}$. So λ' is a k-line, and there are at most $4k$ points below λ (actually exactly, unless $i = \lfloor \log_2 m \rfloor$). We have an edge intersecting ℓ not below λ' but below λ, and the phase starts. The phase ends whenever we reach an edge that intersects ℓ below λ'. Note that for every edge occurring in the phase, one endpoint has to be below λ (since the edge intersects ℓ below λ) and there is an endpoint above λ' (since otherwise, we are already in a new phase).

A few words on what we are heading for. We further split phase i into *strokes*. A stroke starts after we have sampled a point in $\mathrm{below}(\lambda) \cup \mathrm{below}(\lambda')$ (or at the very beginning of the phase) and it finishes after another point in $\mathrm{below}(\lambda) \cup \mathrm{below}(\lambda')$ is chosen (this includes the event that the phase ends); thus, any stroke in the phase terminates with a point in $\mathrm{below}(\lambda) \cup \mathrm{below}(\lambda')$. If N is the number of strokes, then we can write $X := X_i$ as

$$X = Y_1 + Y_2 + \cdots + Y_N$$

where Y_j is the number of iterations of the jth stroke. Note that N itself is a random variable. (For $j > N$ we set $Y_j = 0$.) We will show that

(i) $\mathrm{E}(Y_j \,|\, j \le N) = O(\log n)$ for all j, and

(ii) $\mathrm{E}(N) = O(1)$.

It follows that $\mathrm{E}(X) = O(\log n)$:

$$\begin{aligned} \mathrm{E}(X) &= \sum_{j=1}^{\infty} \overbrace{\mathrm{E}(Y_j \,|\, j \le N)}^{O(\log n)} \Pr(j \le N) \\ &= O(\log n) \sum_{j=1}^{\infty} \Pr(j \le N) \\ &= O(\log n)\,\mathrm{E}(N)\ . \end{aligned} \tag{3}$$

As for the points sampled from $\mathrm{below}(\lambda) \cup \mathrm{below}(\lambda')$ we distinguish several cases depending on where the respective new point pivoted into the current edge lies. We will see that each of these situations is more or less promising in our goal to escape this phase.

Here are the steps in our reasoning: *At any time during the phase*, the following four claims hold.

Claim 1. *The expected number of pivots until we sample a new point in*

$$\text{below}(\lambda)$$

is at most $2\log_2 n$.

Proof. At least one of the two endpoints of the current edge has to be below λ. So in a contiguous subsequence where the new point is always chosen above λ, the other endpoint below stays the same throughout this sequence. We denote this point by q. If we order the points on the other side of ℓ according to their visibility from q, we get almost the situation as in the one-dimensional process described in Figure 6. In fact, there are two differences which can only improve our expectations: We terminate not only in the lowest point but also in $k-1$ other points. In each step we may also sample on q's side of ℓ, in which case we immediately terminate (we have surely sampled below λ). Hence, the expected length of such a subsequence is at most[7] $2\log_2 n$. ◻

Since any new point sampled in below(λ) starts a new stroke, this also establishes our claim (i) from above: the expected number of iterations during a stroke is $O(\log n)$.

Claim 2. *Conditioned on the event that we sample a point in*

$$\text{below}(\lambda) \cup \text{below}(\lambda')\ ,$$

the point will be in

$$\text{below}(\lambda')$$

with probability at least $\frac{1}{5}$.

Proof. Since all edges in this phase intersect ℓ not below λ' it follows: For one side of ℓ, all k points below λ' must also lie below the line through the current edge. That is, *at least* k points below λ' are also below the line through the current edge. On the other hand, *at most* $5k$ points are below λ or λ'. This holds, since $|\text{below}(\lambda)| \leq 4k$, $|\text{below}(\lambda')| = 2k$, and on one side of ℓ, all k points below λ' are also below λ. ◻

[7] Even H_{n-m-1} is true.

Claims 1 and 2 combined assure that we reach a point below λ' within an expected number of at most $10 \log_2 n$ steps.

So what happens after we see such a point p below λ'? Two cases have to be distinguished, depending on whether p is also below λ or not.

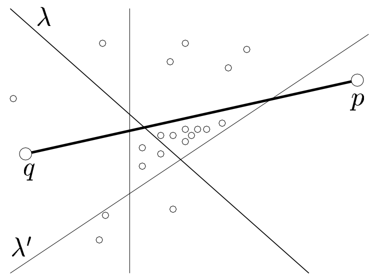


Figure 9: $p \in \text{below}(\lambda') \setminus \text{below}(\lambda)$.

Claim 3. *If an endpoint p of the current edge is in*

$$\text{below}(\lambda') \setminus \text{below}(\lambda) ,$$

then the next point sampled below λ or λ' will be in

$$\text{below}(\lambda) \cap \text{below}(\lambda')$$

with probability at least $\frac{1}{5}$.

Proof. Two relevant conclusions right away (see Figure 9): (i) Since p is not below λ and below λ', while λ' intersects ℓ below λ, the lines λ and λ' must intersect on p's side. (ii) Since p is not below λ, the other endpoint q of the current edge has to be below λ.

Before q can be substituted by a point not below λ, the other endpoint has to be below λ. That is, when we first sample a point below λ or λ', point q is still in the edge. Therefore, the current edge connects q to a point below the edge $\{p, q\}$, above λ'. So the line carrying this edge must intersect λ' on p's side. But then, on q's side, all k points below λ' are also below the then current edge.

Moreover, on q's side, all points below λ' are also below λ (since these lines intersect on the other side). So, summing up, the k points below λ' are both below the current edge and below λ, and they are at disposal, when we sample a point below λ or λ'. The claim follows. $\square$

Claim 4. *If an endpoint p of the current edge is in*

$$\text{below}(\lambda) \cap \text{below}(\lambda') \ ,$$

then the next point sampled below λ or λ' will be in

$$\text{below}(\lambda') \quad \textit{on the side opposite to } p$$

with probability at least $\frac{1}{5}$.

Proof. If p is substituted in a pivot, it must be substituted by a point below λ or λ'. This holds, since on p's side of ℓ, everything below the current edge has to be below λ or λ' (see Figure 10). As a consequence, until the first

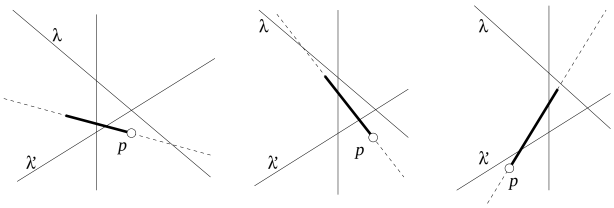


Figure 10: $p \in \text{below}(\lambda') \cap \text{below}(\lambda)$.

pivot with a point below λ or λ', point p is still an endpoint of the edge. But since p is below λ', on the opposite side everything below λ' is also below the current edge. So there are at least k good choices, and at most $5k$ choices of points below λ or λ'. ◻

Claim 4 entails that once we have chosen a point below λ' and λ, then – with probability at least $\frac{1}{5}$ – the next point chosen below λ or λ' will terminate the phase.

To complete the argument, we look at the sequence of points from $\text{below}(\lambda) \cup \text{below}(\lambda')$ that are pivoted into the current edge. Recall that these are exactly the points that terminate the strokes of a phase (except for the last one). If we can show claim (ii) from above, i.e. that the expected length of this sequence is at most some constant c, then the expected length of the whole sequence is at most $2c \log_2 n$ due to Equation 3. Each point in this sequence is classified depending on whether it lies in

$$\begin{array}{ll} \text{Class 0:} & \text{below}(\lambda) \setminus \text{below}(\lambda') \\ \text{Class 1:} & \text{below}(\lambda') \setminus \text{below}(\lambda) \\ \text{Class 2:} & \text{below}(\lambda') \cap \text{below}(\lambda) \end{array}$$

Every point in the sequence considered is in Class 0, 1, or 2. If we have a point in Class 0, the next will be in Class 1 or 2 with probability at least $\frac{1}{5}$ (by Claim 2). If we have a point in Class 1, the next will be in Class 2 with probability at least $\frac{1}{5}$ (by Claim 3). (All of this of course conditioned on the event that a next point exists at all, i.e. the phase hasn't stopped already.) Finally, if we are in Class 2, it is the last point in the sequence with probability at least $\frac{1}{5}$ (by Claim 4).

Now we estimate the expected length of the sequence by the Markov chain[8] depicted in Figure 11, with four states

$$\mathsf{start} = 0, \quad 1, \quad 2, \text{ and } \quad 3 = \mathsf{stop},$$

and the indicated transition probabilities.

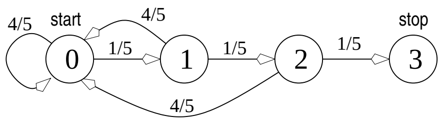


Figure 11: A pessimistic Markov chain.

On one hand, it is easy to calculate that the expected number of steps from start to stop is 155. On the other hand, the chain and our sequence can be coupled so that whenever the chain is in state $s \in \{0, 1, 2, 3\}$, then the corresponding point in the sequence in Class $t \geq s$, or the sequence has ended already. Hence, we have shown that the expected number of pivots in a single phase is bounded by $310 \log_2 n$, and the theorem follows.

Theorem 1 *The expected number of pivots in the process defined in Figure 2 is at most*

$$O((\log n)(1 + \log m)) = O(\log^2 n),$$

where n is the number of points, and m is the smaller of the numbers of points on the two sides of the line.

[8]It simulates a biased coin with success probability $\frac{1}{5}$ and counts the number of experiments until we have three consecutive successes.

The Slow Process. If we consider the process in Figure 2 as an algorithm, it remains to specify how to sample from below(e). It is perhaps worthwhile to mention on the side, that if we can sample efficiently (say in time logarithmic in n), then the process gives a polynomial time algorithm even for exponential size sets. If we sample in the obvious way in $O(n)$ time, then this gives us an $O(n(\log n)^2)$ algorithm. On that we can improve by looking at the alternative slow process[9] in Figure 3. The number of pivots has the same distribution as the process in Figure 2. Here we analyze the number of iterations of the while-loop. To that end, define phases as previously done. Recall that in phase i, the current edge intersects ℓ not below the 2^i-line λ_i.

Claim 5. *For each i, the expected number of iterations in phase i is at most $O(\frac{n}{2^i})$.*

Proof. Divide phase i into strokes as we did it before. That is, a stroke is ended, whenever we sample a point in

$$(\text{below}(\lambda') \cup \text{below}(\lambda)) \cap \text{below}(\varepsilon),$$

(ε the current value of e), or when the phase ends.

In phase i, there are always at least 2^i points from below(λ') that lie below the current edge (on some side of ℓ, all points below λ_i are also below the current edge). That is, at any point, we sample a point resulting in the termination of the stroke with probability at least $\frac{2^i}{n}$. Therefore, the expected number of iterations in a stroke is at most $\frac{n}{2^i}$. The number of strokes is, of course, the same as in the slow process; its expectation is constant. The claim follows. $\square$

Theorem 2 *The expected number of iterations of the process defined in Figure 3 is*

$$\Theta(n)$$

where n is the number of points, unless the starting edge $\{a, b\}$ is already the lowest ℓ-edge.

Proof. $\sum_{i=0}^{\lfloor \log_2 m \rfloor} \frac{n}{2^i} < 2n$ and so the upper bound follows from Claim 5. If $\{a, b\}$ is disjoint from the lowest ℓ-edge, a lower bound of $\frac{3}{2}n$ is obvious, since

[9] The 'below$(e) \neq \emptyset$'-test can be made once in n rounds only, thus causing amortized constant cost. Or, after every pivot, we can go through all points in random order (without replacement) until we find the first point in below(e); if no such point is found, we are done. Compared to the 'pure version', this can only speed up the procedure.

on the average it takes that long until we have sampled both endpoints of the lowest edge at least once. Even if $\{a, b\}$ contains exactly one of the two endpoints of the lowest ℓ-edge, we still need n steps on the average before we meet the other endpoint for the first time. The lower bound follows. ⊒

The coupon collector analysis (cf. [9, Exercise 3.13]) tells us that it takes $\Theta(n \log n)$ iterations until we expect to have sampled each point at least once. The process analyzed finds the lowest ℓ-edge much before all points have been seen at least once.

3 A Lower Bound

We prove a lower bound on the expected number of pivot steps for certain configurations in the process of Figure 2. This is our construction: Let ℓ be the y-axis. We place $4n$ points $p_0, \ldots, p_{2n-1}, q_0, \ldots, q_{2n-1}$ onto the lines $x = \pm 1$ and $x = \pm n$ in such a way that $\{p_k \,|\, k \geq n\} \subseteq \text{below}(\{p_i, q_j\})$ whenever $j < i < n$, and $\{q_k \,|\, k \geq n\} \subseteq \text{below}(\{p_i, q_j\})$, whenever $i < j < n$; cf. Figure 12. Finally, to get general position, we slightly perturb the points.

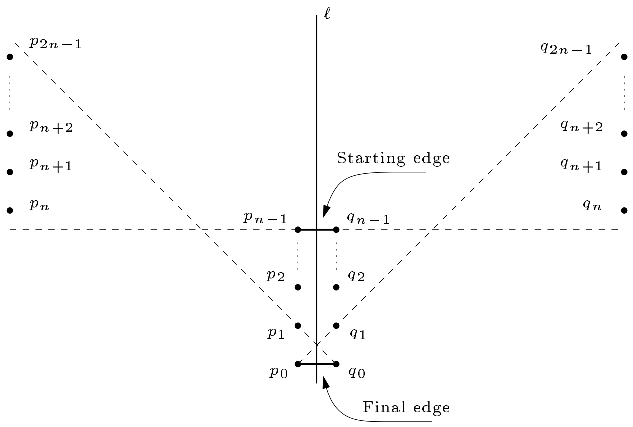


Figure 12: Instance for the lower bound.

Again, we divide the sequence of pivots into distinct phases. We are in phase a as long as $\min(i, j) = a$, where i, j are the indices of the points defining the current edge e.[10]

We define the random variables $X_i, i = 0, \ldots, n-1$, as the number of pivots during the phase i, Z shall denote the total number of pivots under the assumption that we start with the edge $\{p_{n-1}, q_{n-1}\}$. We aim to bound the expected value of Z from below, but a direct approach $\mathrm{E}(Z) = \sum_{i=0}^{n-1} \mathrm{E}(X_i)$ seems not suitable.

Conditioned on the event of choosing the first point in $\{p_0, \ldots, p_i\} \cup \{q_0, \ldots, q_i\}$, p_i and q_i are equally likely chosen with probability $\frac{1}{2(i+1)}$. This implies that

$$\Pr(\text{Phase } i \text{ is entered}) = \begin{cases} \frac{1}{i+1} & \text{for all } i, 0 \leq i < n-1 \\ 1 & \text{for } i = n-1. \end{cases} \tag{4}$$

Thus, denoting the event that phase i is entered by Ph_i, we get

$$\begin{aligned} \mathrm{E}(Z) &= \sum_{i=0}^{n-1} \mathrm{E}(X_i) = \sum_{i=0}^{n-1} \mathrm{E}(X_i \mid \mathrm{Ph}_i) \Pr(\mathrm{Ph}_i) \\ &= \mathrm{E}(X_{n-1} \mid \mathrm{Ph}_{n-1}) + \sum_{i=0}^{n-2} \frac{\mathrm{E}(X_i \mid \mathrm{Ph}_i)}{i+1}. \end{aligned} \tag{5}$$

Recall that a new phase is entered whenever we choose a point whose index sets a new minimum. We say that the phase is on the left (right, respectively) of ℓ whenever its defining minimum point is on the left (right, respectively).

To estimate the expectations of the X_i conditioned on the event that phase i is actually entered we distinguish two cases. If phase i is on the same side as the previous phase, we have to assume the worst: Possibly, it is over after just one flip! But if the minimum switched to the other side with the beginning of phase i then n points ($p_n, \ldots, p_{2n-1}$ or $q_n, \ldots, q_{2n-1}$) are put back into the game.

Using the same argument as for Equation 4, we see that the probability that phase i does not lie on the same side with respect to ℓ as the previous phase is $\frac{1}{2}$.

Let r and s denote the number of points on the left and on the right of ℓ below the current edge; it is easy to see that we are in phase $\min(r, s)$.

The missing piece is to compute $t_{r,s}$, the expected number of flips until we leave the current phase $\min(r, s)$. Clearly, $t_{r,s} = t_{s,r}$, so w.l.o.g. assume

[10] Note that we follow the general trend: Phases count *down* during the process.

$s \leq r$. Then $t_{r,s}$ is monotone in r. For $s < r$ (i.e. in particular, at the beginning of a new phase), we have at least $s+n$ points below the current edge on the side of r.

The simple recursion

$$t_{0,0} := 0\,, \quad t_{r,s} = 1 + \frac{1}{r+s}\sum_{r'=s}^{r-1} t_{r',s} \quad (r+s>0)$$

gives us

$$t_{r,s} = \begin{cases} \mathrm{H}_r & \text{whenever } s=0 \\ \mathrm{H}_{r+s} - \mathrm{H}_{2s} + 1 & \text{otherwise.} \end{cases}$$

Applying these observations to 5 we can therefore deduce that

$$\begin{aligned}
\mathrm{E}(Z) &\geq \sum_{i=0}^{n-2} \frac{\mathrm{E}(X_i \mid \mathrm{Ph}_i)}{i+1} \\
&\geq \sum_{i=0}^{n-2} \frac{\mathrm{E}\left(X_i \mid \mathrm{Ph}_i \wedge \begin{array}{l}\text{side of minimum}\\ \text{changed}\end{array}\right)}{i+1} \Pr\left(\begin{array}{l}\text{side of minimum}\\ \text{changed}\end{array}\right) \\
&\geq \sum_{i=0}^{n-2} \min_{j\geq 0}(t_{i+n+j,i})\frac{1}{2(i+1)} = \sum_{i=0}^{n-2} \frac{t_{i+n,i}}{2(i+1)} \\
&= \frac{\mathrm{H}_n}{2} + \sum_{i=1}^{n-2} \frac{\mathrm{H}_{n+i} - \mathrm{H}_{2i} + 1}{2(i+1)} \geq \frac{1}{2}\left(\sum_{i=0}^{n-2} \frac{\mathrm{H}_n}{i+1} - \sum_{i=1}^{n-2} \frac{\mathrm{H}_{2i}}{i+1}\right) \\
&\geq \frac{1}{2}\left(\mathrm{H}_{n-1}\mathrm{H}_n - \frac{1}{2}\mathrm{H}_{2n-3}^2\right)
\end{aligned}$$

where we used the simple observation $\sum_{i=1}^{n-2} \frac{\mathrm{H}_{2i}}{i+1} \leq \sum_{i=1}^{2n-3} \frac{\mathrm{H}_i}{i} - 1$ as well as the equality

$$\sum_{i=1}^{m} \frac{\mathrm{H}_i}{i} = \frac{1}{2}(\mathrm{H}_m^2 + \sum_{i=1}^{m} \frac{1}{i^2});$$

for a proof see [8, Equation 6.71].

Recalling the well-known estimate $\ln n < \mathrm{H}_n < \ln n + 1$ we conclude

$$E(Z) \geq \frac{1}{4}\ln^2 n + O(\log n). \tag{6}$$

Theorem 3 *There exist instances of one line and n points in general position in the plane such that the expected number of pivot steps for the fast process described in Figure 2 satisfies the following bound:*

$$\mathrm{E}(Z) \geq \frac{1}{4}\ln^2 n + O(\log n).$$

As an encore we would like to describe (without proof) another noteworthy instance for which the same asymptotic lower bound holds: Let ℓ be the y-axis. We place $2n$ points $p_0, \ldots, p_{n-1}, q_0, \ldots, q_{n-1}$ onto the graph of the function $y = f(x) = \log|x|$ such that $x_{p_i} < x_{p_j} < 0 < x_{q_j} < x_{q_i}$ for any $i < j < n$. Furthermore, and most importantly, $\{p_j \,|\, j \neq i\} \subseteq \mathrm{below}(\{p_i, q_i\})$, and $\{q_j \,|\, j \neq i\} \subseteq \mathrm{below}(\{p_{i+1}, q_i\})$, cf. Figure 13.

Note that there exists a sequence of pivots which visits all possible ℓ-edges! That is, there are simple $(n-2)$-polytopes with n facets where we can pivot through all vertices in a monotone fashion (w.r.t. some linear function).

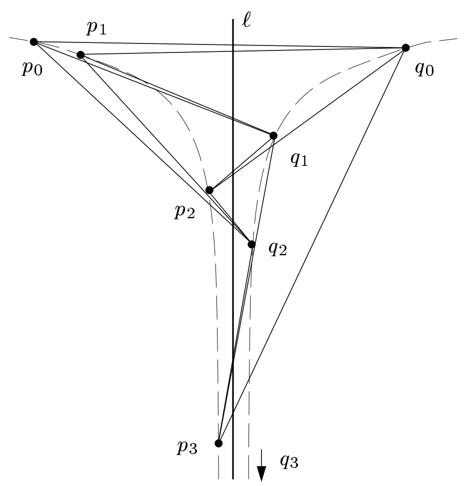


Figure 13: Another instance for the lower bound.

References

[1] Andrei Z. Broder, Martin E. Dyer, Alan M. Frieze, Prabhakar Raghavan, and Eli Upfal. The worst-case running time of the random simplex algorithm is exponential in the height. *Information Processing Letters*, 56(2):79–81, 1995.

[2] Vašek Chvátal. *Linear Programming*. W. H. Freeman, New York, NY, 1983.

[3] George B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey, 1963.

[4] Herbert Edelsbrunner. *Algorithms in Combinatorial Geometry*, volume 10 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Heidelberg, 1987.

[5] Bernd Gärtner, Martin Henk, and Günter M. Ziegler. Randomized simplex algorithms on Klee-Minty cubes. *Combinatorica*, 18:349–372, 1998.

[6] Bernd Gärtner, József Solymosi, Falk Tschirschnitz, Pavel Valtr, and Emo Welzl. One line and n points. *Proc. 33rd Ann. ACM Symp. on the Theory of Computing (STOC)*, 2001.

[7] Bernd Gärtner and Falk Tschirschnitz. Partial chirotopes and the digraph of $(d, d+2)$-polytopes. submitted, 2003.

[8] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, MA, second edition, 1994.

[9] Geoffrey R. Grimmet and David R. Stirzaker. *Probability and Random Processes*. Oxford University Press, New York, NY, second edition, 1992.

[10] Gil Kalai. A subexponential randomized simplex algorithm. In *Proc. 24th Annu. ACM Sympos. Theory Comput.*, pages 475–482, 1992.

[11] Gil Kalai. Linear programming, the simplex algorithm and simple polytopes. *Mathematical Programming Ser. B*, 79:217–233, 1997.

[12] Gil Kalai. Combinatorics with a geometric flavor. In *Geometric and Functional Analysis*, volume II, pages 742–791. Birkhäuser Verlag, Basel, 2000.

[13] D. G. Kelly. Some results on random linear programs. *Methods of Operations Research*, 40:351–355, 1981.

[14] Jiří Matoušek. *Lectures on Discrete Geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, 2002.

[15] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, New York, NY, 1995.

[16] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Prentice Hall, Englewood Cliffs, NJ, 1982.

[17] Falk Tschirschnitz. *LP-related Properties of Polytopes with Few Facets*. PhD thesis, ETH Zürich, 2003.

[18] Emo Welzl. Entering and leaving j-facets. *Discrete & Computational Geometry*, 25(3):351–364, 2001.

APPENDIX: The relationship between Random-Edge and the *Fast Process*

We want to show that there is a one-to-one correspondence between the vertices of a simple $n-2$-polytope with n facets and the ℓ-edges of a suitable point configuration in the plane, with the properties that

(i) the order of the vertices (according to a given generic linear function) matches the order of the corresponding ℓ-edges along ℓ, and

(ii) any pair of adjacent vertices corresponds to a pair of ℓ-edges that share one endpoint.

Theorem 4 below establishes such a correspondence. (Actually, it will more generally relate n-facet polytopes in dimension d to point configurations in $n-d$-space).

Under this correspondence, a Random-Edge pivot step on the polytope becomes a pivot step in the fast process applied to the point configuration corresponding to the polytope, and vice versa. In particular, the expected runtime (number of pivot steps) of Random-Edge, starting at some vertex v, equals the expected length of the fast process, starting from the ℓ-edge corresponding to v.

Consider the following generalisation of the process defined in Figure 2 to arbitrary dimensions:

Let S be a set of n points in general position in $\mathbf{R}^d$ (i.e. no $d+1$ on a common hyperplane), and let ℓ be a vertical line which is disjoint from S and from all intersections of hyperplanes induced by points of S.

The convex hull of a d-tuple $s \in \binom{S}{d}$ defines a $d-1$-simplex, denoted by $\overline{s}$. below(s) denotes the set of points from S that lie below the hyperplane spanned by $\overline{s}$.

We are only concerned with the simplices intersected by ℓ, which we call ℓ-simplices. Note, that given some ℓ-simplex s and a point $p \in$ below(s), there is a unique point $q \in s$, such that the simplex t given by the $d-$tuple $s \cup p \setminus q$ is actually an ℓ-simplex, and we define pivot$(e,p) := t$.

So the *Fast Process* given in Figure 2 can be generalized to the randomized process below. It finds the unique ℓ-simplex s with below$(s) = \emptyset$, given some initial ℓ-simplex $\{p_1, \dots, p_d\}$.

Consider a simple polytope $\mathcal{P}$ of dimension d with $n \geq d+1$ facets, along with a generic linear function $x \to c^T x$, where $c, x \in \mathbf{R}^d$. We will now show that Random-Edge on $\mathcal{P}$ is equivalent to the process given in

$s \leftarrow \{p_1, \ldots, p_d\}$;
while below$(s) \neq \emptyset$ **do**
  $p \leftarrow_{\text{random}}$ below(s);
  $s \leftarrow$ pivot(s, p);

Figure 14: The fast process generalized.

Figure 14 on some configuration of n points and one line in dimension $n-d$.

The problem of maximizing $c^T x$ over $\mathcal{P}$ can be written as

$$\begin{array}{lll} \text{(LP)} & \text{maximize} & c^T x \\ & \text{subject to} & Ax \leq b, \end{array} \tag{7}$$

where A is an $n \times d$-matrix, b an n-vector, and $\{Ax \leq b\}$ a bounded set with some interior point.

Lemma 1 *There exists a strictly positive n-vector $\tilde{y}$ such that $\tilde{y}^T A = 0$.*

Proof. From the fact that $\{Ax \leq b\}$ is bounded, it follows that the linear program

$$\begin{array}{lll} \text{(LP')} & \text{maximize} & \mu_1 + \cdots + \mu_n \\ & \text{subject to} & Ax \leq b - \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \\ & & \mu_i \geq 0, \quad i = 1, \ldots, n \end{array} \tag{8}$$

is bounded; because (LP') is also feasible (set $\mu_i = 0, i = 1, \ldots, n$), it has an optimal solution.

Then, by the LP duality theorem (cf. [2, Chapter 9]), the dual problem

$$\begin{array}{lll} (\text{LP'}^{\Delta}) & \text{minimize} & b^T y \\ & \text{subject to} & A^T y = 0, \\ & & y_i \geq 1, \quad i = 1, \ldots, n \end{array} \tag{9}$$

has an optimal solution, in particular a feasible solution. Any such feasible solution $\tilde{y}$ is a vector with the required properties. $\square$

Assumption 1. For A, b in (7), we can assume the following without loss of generality (in particular, without changing the behavior of RANDOM-EDGE on (7)):

(i) The rows of A sum up to $\mathbf{0}$. This is achieved via Lemma 1, by a suitable scaling of the constraints in (7) by positive multiples.

(ii) Any set of d rows of A is linearly independent. For this, we apply a slight perturbation to A.

(iii) $b > 0$. This is obtained by translating $\mathcal{P}$ in such a way that $\mathbf{0}$ is an interior point.

Lemma 2 *Consider the* $(n+1) \times (d+1)$*-matrix*

$$H := \begin{pmatrix} A & -b \\ c^T & 0 \end{pmatrix}. \tag{10}$$

There exists an $(n-d) \times (n+1)$*-matrix* B *of full row rank with* $BH = \mathbf{0}$*, such that the columns of* B *are in general position when interpreted as points in* $\mathbf{R}^{n-d}$*.*
(We will specify our general position requirements below.)

Proof. The columns of H are vectors in $\mathbf{R}^{n+1}$, spanning a subspace of dimension at most $d+1$. The orthogonal dual of this subspace has therefore dimension at least $n-d$. Choose $n-d$ linearly independent vectors in the orthogonal dual to obtain the rows of B. The prior perturbation of A (we may even perturb H) also lets us choose B in such a way that it assumes any desired general position.[11] $\square$

We need some terminology. For a matrix M and index set I, M_I (M^I, respectively), denotes the submatrix consisting of all rows (columns, respectively) with indices in I. For B as above, let $p_i \in \mathbf{R}^{n-d}$ denote the point corresponding to the i-th column of B, and define

$$\Pi_I := \{p_i, i \in I\}.$$

Theorem 4 *Let* A, b, B *be as above, consider* $I \subseteq [n], |I| = d$ *and set* $\bar{I} := [n] \setminus I$*. Moreover, let* $\tilde{x} := A_I^{-1} b_I$ *(which exists by Assumption 1(ii)) and choose* $\gamma \in \mathbf{R}$*.*

The following two statements are equivalent.

(i) $\tilde{x}$ *is a vertex of* $\mathcal{P}$ *with objective function value* γ*.*

[11] Essentially, we apply the Gale dual transform (cf. [14], [18]) to the vertices of the polar dual of the polytope $\mathcal{P}$ adjoined by the vector c defining the objective function.

(ii) The relative interior of $\mathrm{conv}\Pi_{\bar{I}}$ *intersects the line* $\{tp_{n+1}\}$ *in a unique point, at value* $t = \gamma / \sum_i b_i$.

Proof. Using Lemma 2, we can first argue that

$$\begin{aligned}
\mathbf{0} &= BH \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} \\
&= B^I \underbrace{(A_I | - b_I) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix}}_{\mathbf{0}} + B^{\bar{I}}(A_{\bar{I}} | - b_{\bar{I}}) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} + B^{\{n+1\}}(c^T | 0) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} \\
&= B^{\bar{I}}(A_{\bar{I}} | - b_{\bar{I}}) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} + c^T \tilde{x} p_{n+1}.
\end{aligned}$$

With

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} := -(A \mid -b) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix}$$

it follows that

$$c^T \tilde{x} p_{n+1} = \sum_{i \in \bar{I}} \mu_i p_i, \tag{11}$$

where

$$\begin{aligned}
\sum_{i \in \bar{I}} \mu_i &= (-1, \ldots, -1)(A \mid -b) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} \\
&= (\mathbf{0} \mid \sum_i b_i) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} = \sum_i b_i > 0.
\end{aligned}$$

The second equality is Assumption 1(i), and the final inequality follows from Assumption 1(iii). So, we may scale (11) by $1/\sum_i b_i$, obtaining

$$\frac{c^T \tilde{x}}{\sum_i b_i} p_{n+1} = \sum_{i \in \bar{I}} \lambda_i p_i, \quad \sum_{i \in \bar{I}} \lambda_i = 1, \tag{12}$$

where $\lambda_i = \frac{\mu_i}{\sum_i b_i}$. If statement (i) of the theorem holds, i.e. $\tilde{x}$ is a vertex with $c^T \tilde{x} = \gamma$, then

$$(A_{\bar{I}} | - b_{\bar{I}}) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} < \mathbf{0}$$

by simplicity of $\mathcal{P}$, which implies $\mu_i > 0$ in (11) and $\lambda_i > 0$ in (12). Hence, the point

$$\frac{c^T \tilde{x}}{\sum_i b_i} p_{n+1} = \frac{\gamma}{\sum_i b_i} p_{n+1}$$

is a convex combination of the points in $\Pi_{\bar{I}}$ as described by (12). Statement (ii) follows when we assume that $p_{n+1} \neq \mathbf{0}$ and that the line $\{tp_{n+1}\}$ is disjoint from all affine spaces spanned by less than $n-d$ of the points in $\Pi_{[n]}$. In fact, these are requirements on the 'general position' in Lemma 2.

Now assume that statement (ii) holds and that the set $\Pi_{\bar{I}}$ is affinely independent (our final general position requirement). In this case, there are unique values $\lambda_i > 0$ such that

$$\frac{\gamma}{\sum_i b_i} p_{n+1} = \sum_{i \in \bar{I}} \lambda_i p_i, \quad \sum_{i \in \bar{I}} \lambda_i = 1.$$

Because the line spanned by p_{n+1} intersects $\mathrm{conv}\Pi_{\bar{I}}$ in a single point, there is no other value of γ for which the previous equation can be satisfied. Therefore, there are unique values $\gamma, \mu_i > 0$ with

$$\gamma p_{n+1} = \sum_{i \in \bar{I}} \mu_i p_i, \quad \sum_{i \in \bar{I}} \mu_i = \sum_i b_i. \tag{13}$$

On the other hand, by the above computations, the values

$$\begin{aligned} \gamma &:= c^T \tilde{x}, \\ \mu_i &:= -(A_i| - b_i) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} \end{aligned}$$

satisfy equation (13), so they are the desired unique values. It follows that

$$(A_{\bar{I}}| - b_{\bar{I}}) \begin{pmatrix} \tilde{x} \\ 1 \end{pmatrix} < \mathbf{0},$$

and $\tilde{x}$ is a vertex with $c^T \tilde{x} = \gamma$, proving (i). ◻

It is easy to see that for $d = n-2$ the correspondence $\tilde{x} \leftrightarrow \Pi_{\bar{I}}$ established by the theorem satisfies properties (i) and (ii) from the beginning of this appendix, thus completing the argument.