Random Walks on Polytopes of Constant Corank

Malte Milatz
ETH Zürich
Switzerland
mmilatz@inf.ethz.ch

Abstract

We show that the pivoting process associated with one line and n points in r-dimensional space may need $\Omega(\log^r n)$ steps in expectation as $n \to \infty$. The only cases for which the bound was known previously were for $r \leq 3$. Our lower bound is also valid for the expected number of pivoting steps in the following applications: (1) The Random-Edge simplex algorithm on linear programs with n constraints in $d = n - r$ variables; and (2) the directed random walk on a grid polytope of corank r with n facets.

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

Keywords and phrases polytope, unique sink orientation, grid, random walk

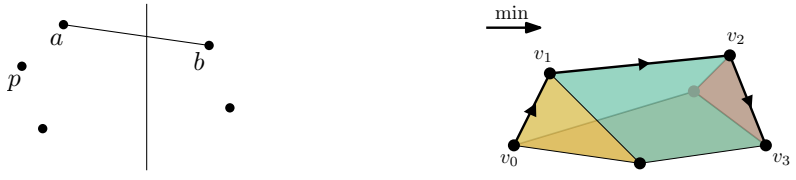
Digital Object Identifier 10.4230/LIPIcs.SoCG.2018.60

Related Version The full version with appendix may be obtained from `https://arxiv.org/abs/1802.09324`.

Acknowledgements I would like to thank Bernd Gärtner, Ahad N. Zehmakan, Jerri Nummenpalo and Alexander Pilz for many useful suggestions and discussions.

1 Introduction

In 2001 Gärtner et al. [6] introduced a random process involving a vertical line and n points in the plane (see Figure 1a) which they called "the fast process" or just "one line and n points". The starting position of this process is a pair $\{a, b\}$ of points that lie on opposite sides of the vertical line. In each step the process picks a point p (a pivot) uniformly at random from the points that lie below the non-vertical line ab. The subsequent position is then the unique pair $\{p, q\}$ with $q \in \{a, b\}$ such that p and q lie again on opposite sides of the vertical line. For example, in Figure 1a the next position would be $\{p, b\}$. The authors of [6] gave matching upper and lower bounds of order $\log^2(n)$ for the expected duration of this process in the plane.



(a) **(b)**

Figure 1 **(a)** An instance of "one line and n points". **(b)** A possible path that the directed random walk might take on the corresponding polytope.

34th International Symposium on Computational Geometry (SoCG 2018).
Editors: Bettina Speckmann and Csaba D. Tóth; Article No. 60; pp. 60:1–60:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The process generalizes naturally to higher dimensions. However, understanding its behavior in any dimension other than 2 has remained a wide open problem, with the notable exception of an $\Omega(\log^3 n)$ lower bound in three dimensions [16]. As the dimension grows, the situation indeed becomes increasingly complicated, and the question has seen no further improvements for the subsequent fifteen years.

The relevance of the generalized process lies in its intimate connection with the Random-Edge simplex algorithm for linear programming, which has already been widely considered before; for example cf. [1, 2, 4, 8, 11]. Random-Edge is naturally formulated in terms of a random walk on the vertex-edge graph of the feasible region (see Figure 1b). To be precise it is a *directed* random walk, because every edge may be used in one prescribed direction only: the direction along which the objective function improves. In each step the directed random walk moves from the current vertex to a vertex chosen uniformly at random from all neighbors with smaller objective value. By means of the so-called *extended Gale transform*, the directed random walk on a d-polytope with n facets translates precisely into the process of "one line and n points" in $\mathbb{R}^r$, where $r := n - d$ denotes the *corank* of the polytope. For lack of space we refer the interested reader to the appendix of [6] for a complete exposition of the extended Gale transform.

The interpretation in terms of polytopes also explains the difficulties that arise when analyzing the process for $r \geq 3$. Namely, it is known that every simple polytope of corank $r = 1$ is a simplex, and every simple polytope of corank $r = 2$ is a product of two simplices; these situations are thus well-understood, classified, and not too complicated. Already for $r = 3$, however, the classification is considerably more involved (cf. chapter 6.3 in [7]), and for $r \geq 4$ no similar classification exists.

The name "corank" is not entirely standard; it has been used e.g. in [14]. Despite its anonymity it plays a prominent role in polytope theory: The *Hirsch conjecture* once stated that the corank might be an upper bound on the diameter of any polytope. Since the Hirsch conjecture, in its strict form, has been disproved by Santos [15], the search for a close connection between these quantities continues. Indeed we believe it fruitful to analyze algorithms for linear programming in a setting where the corank is assumed to be small, i.e., a constant or a slowly growing function of n. A positive result of this type is for example Kalai's Random-Facet algorithm: His upper bound in [10] for the expected number of arithmetic operations becomes polynomial if the corank is taken to be of order $r = O(\log n)$. In contrast, Random-Edge has proved to be notoriously hard to analyze, and tight bounds are rare when we want to understand the behavior of a given simplex algorithm on a complete class of instances. The analysis in [9] for $d = 3$ suggests that Random-Edge can be a good choice in low dimension; the present paper takes the dual viewpoint, fixing the corank rather than the dimension. Note that the mildly exponential lower bound obtained by Friedmann et al. [3] does not pose any restrictions on the corank.

A polytope with n facets and constant corank r has $O(n^r)$ vertices, and this bound is tight; this follows for example from McMullen's theorem [12]. Thus, $O(n^r)$ is a trivial bound for the number of Random-Edge pivoting steps. The known bounds for $r = 2$ suggest that the process outperforms the trivial bound considerably; however, it is not at all clear what can be expected in general. It is conceivable that a bound of $O(\log^r n)$ holds, although we are currently missing the mathematical techniques and insights to prove this. In this paper, we prove that one can at least not do better.

The history of lower-bound constructions for specific LP-solving algorithms shows that it is often considerably easier to prove bounds in an abstract model; this is how *unique sink orientations* (USOs) enter the picture. The same principle applies to the present paper: We

first present a construction in the USO model in Section 2 and strengthen the result to the geometric setting in the rest of the paper. Somewhat atypically, the main ideas underlying the geometric construction are not *entirely* different from the ideas underlying the USO construction.

Our results. We prove that the random process of "one line and n points" in $\mathbb{R}^r$ may take $\Omega(\log^r n)$ steps in expectation. This generalizes previous constructions for the cases $r = 2, 3$ [6, 16]. The exact bound that we obtain is specified in Theorem 13. Using the extended Gale transform, our result can be rephrased in different settings, as in the following theorems. Theorem 1 rephrases our result in the language of linear programs. Theorem 2 uses instead the language of polytopes. The combinatorial type of the polytopes that we construct in this way is rather special: they are *grid polytopes*, i.e., Cartesian products of simplices.

▶ **Theorem 1.** *Let $r \in \mathbb{N}_0$ be a fixed parameter. There are linear programs in d variables with $n = d + r$ constraints on which the* Random-Edge *algorithm needs $\Omega(\log^r n)$ pivoting steps in expectation as $n \to \infty$.*

▶ **Theorem 2.** *Let $r \in \mathbb{N}_0$. There are grid d-polytopes with $n = d + r$ facets on which the directed random walk (with the direction specified by a linear function) needs $\Omega(\log^r n)$ steps in expectation as $n \to \infty$.*

2 Prelude: walks on grids

In this section we prove a lower bound of order $\log^r(n)$ for the expected duration of a random walk on a certain class of directed graphs, namely *unique sink orientations of grids*. The construction does not involve any geometry and should be simpler to understand than the point-set construction in Section 4.

Given a directed (multi-)graph G and a vertex v_0 of G, the *directed random walk* is the random process $v_0, v_1, v_2, \ldots$ described as follows: If the current position is v_i, choose one of the outgoing edges at v_i uniformly at random, and let v_{i+1} be the other endpoint of that edge. The process terminates when (and if) it reaches a sink. The random variable

$$T(G, v_0) = \min\{t : v_t \text{ is a sink}\}$$

will denote the duration of the directed random walk on G starting in v_0. We will abbreviate $T(G) := T(G, v_0)$ for a starting position v_0 chosen uniformly at random from the set of vertices.

Following the terminology in [5], a *grid* is a Cartesian product of a (finite) number of (finite) complete graphs $G_1, \ldots, G_r$. Explicitly, the vertex set of this graph is $V(G_1) \times \cdots \times V(G_r)$, and two vertices $(u_1, \ldots, u_r)$ and $(v_1, \ldots, v_r)$ are joined by an edge if and only if they differ in exactly one coordinate. We could have defined grids equivalently as the vertex-edge graphs of grid polytopes. The number r is usually called the *dimension* of the grid (this is *not* the dimension of the underlying grid polytope!); and its *size* is the number $|V(G_1)|+\cdots+|V(G_r)|$.

A *subgrid* of a grid is an induced subgraph on a set of the form $U_1 \times \cdots \times U_r$ with $U_i \subseteq V(G_i)$. Finally, a *unique sink orientation* of a grid is an orientation of the edges of the grid with the property that every subgrid possesses a unique sink. An example is shown in Figure 2a.

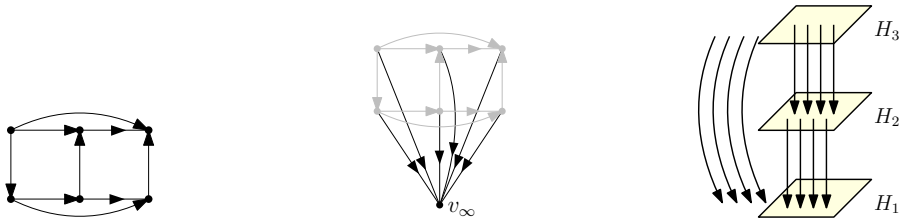


Figure 2 **(a)** A grid G of dimension $r = 2$ and size $n = 5$. **(b)** The augmented (multi-)graph G^Δ, here for $\Delta = 1$. **(c)** A schematic depiction of the grid constructed in the proof of Theorem 3.

▶ **Theorem 3.** *Let $r \in \mathbb{N}_0$ and $m \in \mathbb{N}$. There is an r-dimensional grid G of size $n := rm$, endowed with an acyclic unique sink orientation, such that a directed random walk on the grid, starting at a random position, takes at least*

$$\mathrm{E}\big[T(G)\big] \geq \frac{1}{r!} \ln^r (m+1) - 1 \tag{1}$$

steps in expectation.

We remark that the theorem is meaningful primarily for a fixed dimension r and with the grid size tending to infinity. In this setting it says that the number of pivoting steps can be of order $\Omega(\log^r n)$.

Proof. We can choose the grid to be $G = G_1 \times \cdots \times G_r$ with $G_1 = \cdots = G_r = K_m$. We construct a unique sink orientation on G by induction on r. For $r = 0$, the graph consists of a single vertex, so we need not orient any edges.

For $r \geq 1$, we first choose a permutation of $V(G_r)$ uniformly at random from the set of all permutations, and we label the vertices according to the chosen permutation, as in $V(G_r) = \{1, \ldots, m\}$. Now we partition the grid into "hyperplanes"

$$H_i = G_1 \times \cdots \times G_{r-1} \times \{i\} \quad (i = 1, \ldots, m).$$

Each hyperplane H_i is a subgrid of dimension $r-1$, so we can inductively assign an orientation to it. We do this for each hyperplane independently (i.e., all random permutations used throughout the construction are chosen independently). The only edges that we still need to orient are those between vertices from two distinct hyperplanes. We orient those according to our chosen permutation of $V(G_r)$. Explicitly, the edge from a vertex $(u_1, \ldots, u_{r-1}, i)$ to $(u_1, \ldots, u_{r-1}, j)$ is directed forwards if and only if $i > j$. See Figure 2c for an illustration.

It is easy to verify that this defines an acyclic unique sink orientation of the grid. We will now analyze the duration of the random walk from a random starting position. Concerning the starting position, here is a key observation: Due to the random permutations involved in the construction of the grid orientation, it amounts to the same random process whether we start the walk in a random position of the grid, or whether we start in any fixed given position. Consequently, if the random walk visits one of the hyperplanes H_i, we can relate the behavior within H_i to the $(r-1)$-dimensional construction, for which we have a lower bound by induction. However, there *is* a difference between H_i and the $(r-1)$-dimensional construction: namely, every vertex of H_i has $i-1$ additional outgoing edges by which the walk may leave H_i at any moment. To account for these, we make use of the following "augmented" multigraph; see Figure 2b for an example.

▶ **Definition 4.** Given any (multi-)graph Γ and a parameter $\Delta \in \mathbb{N}_0$, we define an *augmented multigraph* Γ^Δ as follows. We add a new, special, vertex v_∞ to the vertex set of Γ. Furthermore we add Δ many edges $\overrightarrow{vv_\infty}$ for every $v \in V(\Gamma)$. If $\Delta = 0$ then we add one additional edge $\overrightarrow{sv_\infty}$ for every sink s of Γ; this way we ensure that v_∞ is the only sink of Γ^Δ.

▶ **Lemma 5.** *Let $\Delta \in \mathbb{N}_0$. Then the expected duration of the directed random walk on the augmented construction G^Δ, starting from a random position (and ending in v_∞), satisfies the bound*

$$\mathrm{E}\big[T(G^\Delta)\big] \geq \frac{1}{r!}\left(\ln\left(m+\Delta+1\right) - \ln(\Delta+1)\right)^r.$$

Proof of the lemma. We proceed by induction on $r \geq 0$. If $r = 0$ then G consists of a single vertex and there is nothing to prove; so let $r \geq 1$. For $i \in \{1,\dots,m\}$, let T_i denote the number of positions that the walk visits on H_i, so that the total duration of the walk is given by

$$T(G^\Delta) = T_1 + \cdots + T_m. \tag{2}$$

Furthermore, let $\mathcal{E}_i$ denote the event that the random walk visits at least one vertex in H_i $(i = 1,\dots,m)$. We claim

$$\Pr\big[\mathcal{E}_i\big] \geq \frac{1}{\Delta+i}. \tag{3}$$

To this end we consider the hitting time

$$\tau_i := \min\big\{t \;:\; v_t \in \{v_\infty\} \cup H_1 \cup \cdots \cup H_i\big\}$$

where, as before, $v_t \in V(G^\Delta)$ denotes the position that the random walk visits at time t $(t = 0, 1, 2, \dots)$. Note that the hyperplanes H_i are visited in decreasing order; so either the walk visits H_i at time τ_i, or not at all. Hence, $\mathcal{E}_i$ equals the event $\{v_{\tau_i} \in H_i\}$. We now calculate the probability of this event by conditioning on $\tau_i \geq 1$.

Case 1: $\tau_i \geq 1$. Since $v_{\tau_i - 1}$ has Δ outgoing edges to v_∞ and one outgoing edge to each of the hyperplanes $H_1,\dots,H_i$, and since the random walk is equally likely to move along any of these $\Delta + i$ edges, we obtain

$$\Pr\big[\mathcal{E}_i \,\big|\, \tau_i \geq 1\big] = \Pr\big[v_{\tau_i} \in H_i \,\big|\, \tau_i \geq 1\big] = \frac{1}{\Delta+i}. \tag{4}$$

Case 2: $\tau_i = 0$. Here we need to look at v_0, which (conditioned on $\tau_i = 0$) is a vertex taken uniformly at random from the set $H_1 \cup \cdots \cup H_i$. Since the hyperplanes $H_1,\dots,H_i$ are all of equal cardinality, we obtain

$$\Pr\big[\mathcal{E}_i \,\big|\, \tau_i = 0\big] = \Pr\big[v_0 \in H_i \,\big|\, \tau_i = 0\big] = \frac{1}{i} \geq \frac{1}{\Delta+i}. \tag{5}$$

The claim (3) follows by combining (4) and (5). Now, it is easy to see that $T_i|\mathcal{E}_i$ has the same distribution as $T((H_i)^{\Delta+i-1})$, so that we obtain

$$\begin{aligned}\mathrm{E}\big[T_i\big] &= \Pr\big[\mathcal{E}_i\big] \cdot \mathrm{E}\big[T_i \,\big|\, \mathcal{E}_i\big] \\ &\geq \frac{1}{\Delta+i} \cdot \mathrm{E}\big[T((H_i)^{\Delta+i-1})\big]\end{aligned}$$

$$\geq \frac{1}{\Delta+i} \cdot \frac{1}{(r-1)!} \big(\ln(m+\Delta+i) - \ln(\Delta+i)\big)^{r-1}$$

where the last step was using the induction hypothesis. With (2) we obtain

$$\begin{aligned}
\mathrm{E}[T] &= \sum_{i=1}^{m} \mathrm{E}[T_i] \\
&\geq \sum_{i=1}^{m} \frac{1}{\Delta+i} \cdot \frac{1}{(r-1)!} \big(\ln(m+\Delta+i) - \ln(\Delta+i)\big)^{r-1} \\
&\geq \sum_{i=1}^{m} \frac{1}{\Delta+i} \cdot \frac{1}{(r-1)!} \big(\ln(m+\Delta+1) - \ln(\Delta+i)\big)^{r-1} \\
&\geq \int_{1}^{m+1} \frac{1}{\Delta+x} \cdot \frac{1}{(r-1)!} \big(\ln(m+\Delta+1) - \ln(\Delta+x)\big)^{r-1} \,\mathrm{d}x \\
&= \Big[-\frac{1}{r!} \big(\ln(m+\Delta+1) - \ln(\Delta+x)\big)^{r} \Big]_{x=1}^{m+1} \\
&= \frac{1}{r!} \big(\ln(m+\Delta+1) - \ln(\Delta+1)\big)^{r}
\end{aligned}$$

which concludes the proof of the lemma. In order to deduce the theorem, we choose $\Delta = 0$ to obtain a random orientation (i.e., a probability distribution of orientations) of G^0 such that

$$\mathrm{E}[T(G^0)] \geq \frac{1}{r!} \ln^r(m+1). \tag{6}$$

A directed random walk on G corresponds to a random walk on G^0, except that the latter does one additional step in the end (from the sink of G to the extra vertex v_∞). Thus we need to subtract 1 from the bound (6) to obtain the desired bound (1). We are left only to observe that there must then also exist at least one concrete (not random) choice of orientation G that satisfies this bound. This concludes the proof of Theorem 3. ◀

In Theorem 3 we chose n to be a multiple of r. For other values of n we can still deduce essentially the same bound, as in the following corollary.

▶ **Corollary 6.** *For all $r, n \in \mathbb{N}$ with $n > r$ there is an r-dimensional acyclic unique sink orientation of a grid $\hat{G}$ of size n such that $\mathrm{E}[T(\hat{G})] > \frac{1}{r!} \ln^r\left(\frac{n}{r}\right) - 1$.*

Proof. If n is divisible by r, then the corollary is an immediate consequence of Theorem 3. Assume that n is not divisble by r, and let $m = \lfloor \frac{n}{r} \rfloor$. We take our construction G from the proof of Theorem 3 for the size $n' := rm$ and embed it into a grid $\hat{G}$ of size n. We can do this in such a fashion that all the edges between G and $\hat{G} \setminus G$ point into G. In this way we obtain an acyclic unique sink orientation on $\hat{G}$ with $\mathrm{E}[T(\hat{G})] \geq \mathrm{E}[T(G)]$, which yields the corollary using (1) with $m > \frac{n}{r} - 1$. ◀

3 One line and n points

Here we describe the geometric setting in which we prove our main theorem. We will assume that we are given a set of n points $A \subseteq \mathbb{R}^r$ and a non-zero vector $u \in \mathbb{R}^r$. Its linear span $\mathbb{R}u$ is a line: the "one line" or *requirement line* featured in the heading of this section.

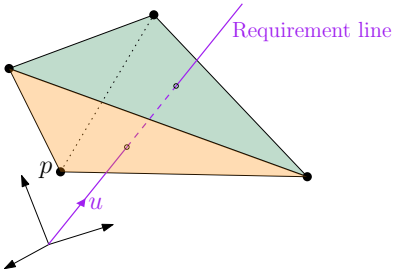


Figure 3 A tetrahedron in $\mathbb{R}^3$. Its two front facets, the green and orange triangles, are both a pierced simplex. If S denotes the green triangle at the top, then the *simplex obtained by pivoting at S with p* is the orange triangle at the bottom.

Pierced simplices and general position. We call a set $S \subseteq A$ *pierced* or, more exactly, *u-pierced* if the convex hull $\mathrm{conv}(S)$ intersects the requirement line $\mathbb{R}u$. If in addition $|S| = r$, then S is a *pierced simplex*. Some readers might find it more natural to reserve the term "simplex" for the set $\mathrm{conv}(S)$ instead of S; we will always take care to distinguish between the two whenever the distinction is important. A pierced simplex S is *non-degenerate* if (I) S is affinely independent, (II) no proper subset of S is a pierced set, and (III) the affine hyperplane spanned by S and the requirement line are not parallel. — Within this paper, when we say that (A, u) is *in general position*, we merely mean that every u-pierced simplex $S \subseteq A$ is non-degenerate.

Below and above. Consider the affine span of a non-degenerate pierced simplex $S \subseteq A$: it is a hyperplane in $\mathbb{R}^r$ that is not parallel to the vector u. The direction of u thus determines an orientation (a positive and negative side) of this hyperplane. For any point $x \in A$, we will say that x lies *(strictly) below* S if it lies on the (strictly) negative side of the hyperplane. The word "above" is understood similarly.

Pivoting steps. Given (A, u) in general position, a pierced simplex $S \subseteq A$ and a point $p \in A$ strictly below S, we define a new pierced simplex $S' \subseteq A$ which we call the *simplex obtained by pivoting at S with p*. To this end consider the set $S \cup \{p\}$: It is an r-dimensional simplex, and the boundary of $\mathrm{conv}(S \cup \{p\})$ is pierced by the line $\mathbb{R}u$ exactly twice: once in the facet S, and once in another facet which we take to be S'. See Figure 3 for an example. We note that

- S' is by general position uniquely determined,
- S' is a subset of $S \cup \{p\} \subseteq A$, and
- S' is a pierced simplex.

The random process. Given a finite set $A \subseteq \mathbb{R}^r$ of n points, a non-zero vector $u \in \mathbb{R}^r$, and a u-pierced simplex $S_0 \subseteq A$, we define the following random process, denoted $\mathcal{R}(A, u, S_0)$. We will keep on assuming that (A, u) is in general position.

The states (or positions) of the process are pierced simplices, and S_0 is the starting position. The consecutive positions $S_1, S_2, \ldots$ are obtained as follows. If the current position is S_i, let p_i be a point (the ith "pivot") chosen uniformly at random from the set of points from A that lie strictly below S_i. (If there are no such points, then the random process terminates at this stage.) Now define S_{i+1} to be the simplex obtained by pivoting at S_i with p_i. — Our main theorem in Section 5 states that the expected number of steps until the process terminates can be of order $\log^r(n)$.

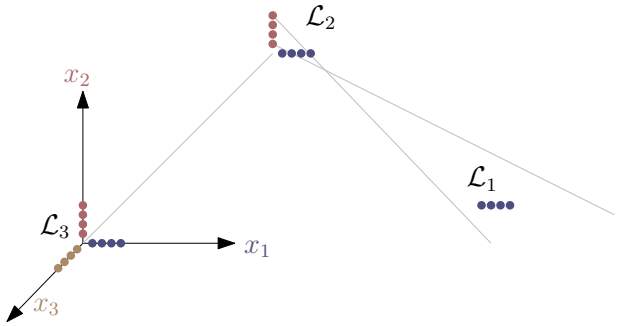


Figure 4 The constructed point set for $r = 3$, $m = 4$. The sketch is not true to scale. All off-axis points lie in the plane $x_3 = -64$.

4 Construction

Here we construct the point set that underlies the proof of our main theorem (Theorem 13 in Section 5).

Points, colors, layers, and phases. For all $r, m \in \mathbb{N}$ we define our point set $A(r, m) \subseteq \mathbb{R}^r$ as follows; a sketch is shown in Figure 4. We use the notation $\mathbf{0}_r = (0, \ldots, 0)$ for the all-zeros vector in $\mathbb{R}^r$. We let

$$A(r, m) := \{\, a_{i,j,k} \,:\, i, j \in [r],\ i \le j,\ k \in [m] \,\}$$

where

$$a_{i,j,k} := \begin{pmatrix} \mathbf{0}_{i-1} \\ \left(m^3 + m^5 + \cdots + m^{2r-2j+1}\right) + (r-j)m + k \\ \mathbf{0}_{j-i} \\ -m^{2r-2j+1} \\ -m^{2r-2j-1} \\ \vdots \\ -m^5 \\ -m^3 \end{pmatrix}.$$

In particular, for $j = r$, we have $a_{i,r,k} = k\mathbf{e}_i$, where $\mathbf{e}_i$ denotes the ith standard unit vector in $\mathbb{R}^r$. We call the indices i, j, k the *color*, the *layer*, and the *phase* of a point, respectively. Sometimes we will need a notational shorthand for colors and layers, so we define $\mathcal{C}_i \subseteq A(r, m)$ to denote the set of points of color i, and $\mathcal{L}_j \subseteq A(r, m)$ to denote the set of points from layer j. So defined, $\mathcal{L}_j$ consists of jm points, and our point set consists of $n = \binom{r+1}{2} \cdot m$ points overall.

We will fix u to denote the all-ones vector, $u = \mathbf{1}_r$, so that a set $S \subseteq A(r, m)$ is pierced if and only if its convex hull intersects the diagonal line $\mathbb{R}\mathbf{1}_r$. The rest of this section is devoted to a number of lemmas concerning the relevant structure of our construction. We begin by identifying the pierced subsets.

▶ **Lemma 7.** *Let $S \subseteq A(r, m)$ be a pierced subset. Then S contains a point of color i, for all $i \in [r]$.*

The next lemma shows the converse of Lemma 7, implying that the set of pierced simplices can be identified with the set $\mathcal{C}_1 \times \cdots \times \mathcal{C}_r$. More to the point, this implies that the extended Gale transform of our point set defines a grid polytope.

▶ **Lemma 8.** *Let $S \subseteq A(r,m)$ and assume that S contains a point from each color class $\mathcal{C}_1, \ldots, \mathcal{C}_r$. Then S is a pierced subset. Furthermore, for each $i \in [r]$, $\operatorname{conv}(S)$ intersects the ith coordinate axis in some point $t\mathbf{e}_i$ with $t > 0$.*

▶ **Lemma 9.** *$(A, \mathbf{1}_r)$ is in general position; that is, every pierced simplex $S \subseteq A(r,m)$ is non-degenerate.*

Lemma 9 above assures that the random process associated with our construction is well-defined. The next lemma states that, as our random process evolves, the intersection value t of the current position with the ith coordinate axis ($i = 1, \ldots, r$) is monotonically decreasing with time and, thus, can serve as a measure of progress. This is of course by no means true for an arbitrary point set, but it holds in the case of our construction. Next, Lemma 11 states a simple condition from which to tell whether the points from the layer $\mathcal{L}_{r-1}$ lie above or below the current position; this condition is immediately relevant for the analysis of the random process.

▶ **Lemma 10.** *Let $i \in [r]$. Let $S \subseteq A(r,m)$ be a pierced simplex, let $p \in A$ be a point strictly below S, and let S' denote the pierced simplex obtained by pivoting at S with p. Let $t_i\mathbf{e}_i$ and $t'_i\mathbf{e}_i$ denote the intersection of $\operatorname{conv}(S)$ (respectively, $\operatorname{conv}(S')$) with the ith coordinate axis, as in Lemma 8. Then we have $t'_i \leq t_i$ for all i.*

▶ **Lemma 11.** *Let $r \geq 2$, $m \geq 2$, let $S \subseteq A(r,m)$ be a pierced simplex, and let $t_i\mathbf{e}_i$ denote the intersections of $\operatorname{conv}(S)$ with the ith coordinate axis as in Lemma 8 ($i = 1, \ldots, r$).*

(a) *If, for some i, $t_i \leq t_r$, then all points from $\mathcal{C}_i \cap \mathcal{L}_{r-1}$ lie strictly* above *S.*

(b) *If, for some i, $t_i \geq t_r + 1$, then all points from $\mathcal{C}_i \cap \mathcal{L}_{r-1}$ lie strictly* below *S.*

(c) *If i is such that $t_i = \min\{t_1, \ldots, t_r\}$, then all points from $\mathcal{C}_i \setminus \mathcal{L}_r$ lie strictly* above *S. In particular we then have $t_i\mathbf{e}_i \in S$.*

The last lemma in this section states that taking the point set $A(r+1,m)$ and removing the outermost layer $\mathcal{L}_{r+1}$ yields a point set that is equivalent, for our purposes, to the set $A(r,m)$. This observation is key to the inductive approach followed in section 5. Actually, the statement is a bit more general: The lemma starts from the set $A(R,m)$ for any $R > r$ and then removes all higher layers $\mathcal{L}_{r+1}, \ldots, \mathcal{L}_R$.

▶ **Lemma 12.** *Assume $m \geq 3$. For $R > r$, let*

$$B := \{(x_1, \cdots, x_r) : x \in A(R,m),\ x \in \mathcal{L}_1 \cup \cdots \cup \mathcal{L}_r\}.$$

Then Lemmas 7 to 11 are also valid for the point set B in place of $A(r,m)$.

5 Analysis

The goal of this section is to prove the main theorem of this paper, concerned with the random process $\mathcal{R}_{r,m} = \mathcal{R}(A, u, S_0)$ associated with the point set $A := A(r,m)$, the all-ones vector $u := \mathbf{1}_r$, and the starting position $S_0 := \{m\mathbf{e}_1, \ldots, m\mathbf{e}_r\}$:

▶ **Theorem 13.** *The expected number of steps performed by the random process $\mathcal{R}_{r,m}$ is at least*

$$\frac{1}{r!^3}\left(\ln\left(m + \tbinom{r}{2} + \Delta\right) - \ln\left(1 + \tbinom{r}{2} + \Delta\right)\right)^r = \Omega(\log^r m).$$

In terms of the number of points, n, the bound can be written in the form $\Omega(\log^r n)$.

The augmented process $\mathcal{R}^{\Delta}_{r,m}$. In order to make an inductive proof possible, we will make use of an "augmented" pivoting process, in analogy to the "augmented graph" that we used in the proof of Theorem 3. Given a number $\Delta \geq 0$, the *augmented process* $\mathcal{R}^{\Delta}_{r,m}$ is defined as follows.

- The starting position is chosen by an adversary, in the following way. The adversary chooses one new point $\alpha_i \mathbf{e}_i$ on each axis, subject to the constraint $\alpha_i \geq m$. These points are added to the point set $A = A(r,m)$ to obtain an augmented point set

  $$A' := A \cup \{\alpha_i \mathbf{e}_i \,:\, i \in [r]\},$$

  and the starting position is now chosen as $S_0 := \{\alpha_1 \mathbf{e}_1, \ldots, \alpha_r \mathbf{e}_r\}$. We fit the new points into our terminology of colors, layers and phases by saying that $\alpha_i \mathbf{e}_i$ has color i, layer r and phase $m+1$; and we remark that the lemmas in Section 3 still hold for the augmented point set.
- The positions $S_0, S_1, \ldots$ of the augmented process are pierced simplices of A', except that we also introduce a new, special, position S_∞. (S_∞ is just a formal symbol; it is not represented by any simplex.) This will be the terminal position.
- If we are currently at position S_t, then the next position S_{t+1} is obtained as follows: Let $\mathrm{below}(S_t) \subseteq A'$ denote the set of points that lie strictly below S_t. We draw a pivot element p_{t+1} from the set $\mathrm{below}(S_t) \cup \{\infty\}$ according to the distribution

  $$\Pr\big[p_{t+1} = x\big] = \begin{cases} \frac{1}{|\mathrm{below}(S_t)|+\Delta} & \text{for } x \in \mathrm{below}(S_t), \\ \frac{\Delta}{|\mathrm{below}(S_t)|+\Delta} & \text{for } x = \infty. \end{cases}$$

  If $p_{t+1} = \infty$, then $S_{t+1} = S_\infty$, and the process terminates. Otherwise we perform a standard pivoting step at S_t with p_{t+1}. (Edge case: If $\Delta = 0$ and $\mathrm{below}(S_t) = \emptyset$, then we always pick $p_{t+1} = \infty$.)

Note that, despite its name, the "augmented" process typically terminates earlier than the non-augmented process: the larger the parameter Δ is, the sooner! For $\Delta = 0$ the augmented process behaves like the original, non-augmented process — except for the modified starting position and one additional final pivoting step towards the terminal position S_∞.

The phase of a pierced simplex. We define the *phase* of a pierced simplex $S \subseteq A'$ as

$$\mathrm{phase}(S) := \min\{\mathrm{phase}(p) \,:\, p \in S \text{ with } \mathrm{layer}(p) = r\}.$$

The minimum is well-defined because S contains at least one point from the layer $\mathcal{L}_r$: Indeed Lemma 7 tells us that S contains a point from $\mathcal{C}_r$, which is a subset of $\mathcal{L}_r$. For consistency we also define $\mathrm{phase}(S_\infty) := 0$. — Note that we are overloading the term "phase", because we have defined the phase of a point earlier.

We remark that the phase of a pierced simplex can be equivalently written as $\mathrm{phase}(S) = \min\{t_1, \ldots, t_r\}$, where $t_i \mathbf{e}_i$ denotes the intersection of $\mathrm{conv}(S)$ with the ith coordinate axis; this follows from Lemma 11(c). Using this observation, the possible choices for pivots at the time of a phase change are easily identified, and from this information we may read off the probability that a particular phase is visited. The following lemma summarizes the result of this observation.

▶ **Lemma 14** (The phases visited by the augmented process). *Let $\sigma_1 < \sigma_2 < \cdots < \sigma_N$ denote the times at which a phase change occurs in the augmented random process, and let $(\phi_i)_{0 \leq i \leq N}$ denote the phases that are visited, i.e., $\phi_0 = m+1$, and $\phi_i = \mathrm{phase}(S_{\sigma_i})$ $(1 \leq i \leq N)$. Then we have, for $i \geq 1$:*

(a) *The distribution of ϕ_i is given by*

$$\Pr\big[\phi_i = x \,\big|\, \phi_{i-1}\big] = \begin{cases} \frac{r}{r(\phi_{i-1}-1)+\Delta} & \text{for } x \in [\phi_{i-1}-1], \\ \frac{\Delta}{r(\phi_{i-1}-1)+\Delta} & \text{for } x = 0. \end{cases}$$

(b) *If $\phi_i > 0$, then the color of the pivot at time σ_i is a u.a.r. element of $[r]$.*

Consider one of the phases ϕ_i that are visited by the augmented random process. We want to bound the *duration* of the phase ϕ_i, i.e. the number $\sigma_{i+1} - \sigma_i$, from below. In general this duration could be very short, so we introduce a suitable notion of a *good phase*. The definition will guarantee that, when entering a good phase, all points of the layer $\mathcal{L}_{r-1}$ will lie strictly below the current position; and this property will in turn make it possible to derive a lower bound on the duration of a good phase inductively.

▶ **Definition 15** (good phases)**.** Let $k \in [m]$. We say that k is a *good phase* of the augmented process if, using the notation from Lemma 14,

(i) k is visited, so that $k = \phi_j$ for some j,
(ii) the pivot at time σ_j has color r, and
(iii) the position S_{σ_j} does not contain any point from the layer $\mathcal{L}_{r-1}$.

Let the reader be warned that the above definition is weaker than one might think at first: The only points that we take directly into account are those from the two outermost layers $\mathcal{L}_r$ and $\mathcal{L}_{r-1}$. In particular we allow S_{σ_j} to contain points from other layers. When reading on, it is useful to keep in mind one consequence of Lemma 11(c): The phase of the current position can change only when pivoting a point from the layer $\mathcal{L}_r$, i.e., a point that lies on one of the coordinate axes. Consequently, pivots in lower layers can largely be ignored in our analysis.

▶ **Lemma 16.** *Let $1 \le k \le m-1$. Then phase k is a good phase with probability at least $\frac{1}{r(\Delta+kr)}$.*

Proof. Let j be the (random) largest index such that $\phi_{j-1} > k$. Then the probability of (i) equals $\Pr\big[\phi_j = k\big]$, and using Lemma 14(a) we compute this probability to be $\frac{r}{rk+\Delta}$. Given (i), Lemma 14(b) tells us that the probability of (ii) equals $1/r$.

Assume that (i) and (ii) hold. It remains to show that, in this case, (iii) holds with probability at least $1/r$. We consider the position S_{σ_j-1} one time step before entering phase k. Using the same notation as in Lemma 11, let $t_i\mathbf{e}_i$ denote the intersections of $\operatorname{conv}(S_{\sigma_j-1})$ with the ith coordinate axis ($i = 1, \dots, r$). Note that $t_i > k$ for all i, because phase k has not been entered yet at this time.

By Lemma 11 it is sufficient to give a bound for the event

$$t_i \le t_r \text{ for all } i = 1, \dots, r-1. \tag{7}$$

To this end, let τ denote the time that the point $t_r\mathbf{e}_r$ is pivoted, so that S_τ is the first position to include the point $t_r\mathbf{e}_r$. Note that t_r does not change in between time τ and time σ_j; thus, if property (7) already holds at time τ, then by monotonicity (Lemma 10) it will still hold at time S_{σ_j-1}. So assume that at time τ property (7) does not yet hold, so that there are some "bad" indices i with $t_i > t_r$. Let $I \subseteq [r-1]$ denote the set of such bad indices, and let $\tau_1 > \tau$ be the first time that another point of layer r with phase $\le t_r$ and color contained in $I \cup \{r\}$ is pivoted. With probability at least $\frac{|I|}{|I|+1}$, the color of this pivot

is contained in I, in which case the number of bad indices is reduced by 1. Iterating this argument, the number of bad indices will be reduced down to zero with probability at least

$$\frac{|I|}{|I|+1} \cdot \frac{|I|-1}{|I|} \cdots \frac{1}{2} = \frac{1}{|I|+1} \geq \frac{1}{r},$$

as desired. ◀

When the augmented process enters a good phase k, then all the points of the layer $\mathcal{L}_{r-1}$ lie strictly below the current position. We restrict our attention to the hyperplane $x_r = -m^3$ that contains all the points from $A(r,m) \setminus \mathcal{L}_r$: Due to Lemma 12, the augmented process within this hyperplane behaves like the augmented process on the lower-dimensional construction $A(r-1,m)$. However, the process might at any point pivot one of the points $\{p \in \mathcal{L}_r : \text{phase}(p) < k\}$, and as soon as this happens, the good phase k already ends. We can account for this by adjusting the parameter Δ and we obtain the following lemma.

▶ **Lemma 17.** *For every $1 \leq k \leq m-1$, if phase k is visited and if it is a good phase, then its expected duration is bounded from below by the best-case*[1] *expected duration of the process $\mathcal{R}^{\Delta+(k-1)r}_{r-1,m}$.*

▶ **Theorem 18.** *Let $t^{\Delta}_{r,m}$ denote the best-case expected duration of the augmented process $\mathcal{R}^{\Delta}_{r,m}$. Then we have*

$$t^{\Delta}_{r,m} \geq \frac{1}{r!^3} \cdot \Big(\ln(m + \tbinom{r}{2} + \Delta) - \ln(1 + \tbinom{r}{2} + \Delta)\Big)^r.$$

Proof. By induction on r. For $r = 1$ the statement is easy to verify; let now $r \geq 2$. Combining Lemmas 16 and 17 we obtain

$$t^{\Delta}_{r,m} \geq \sum_{k=1}^{m-1} \frac{1}{r(kr+\Delta)} \cdot t^{\Delta+kr-1}_{r-1,m}. \tag{8}$$

The induction hypothesis gives, for $1 \leq k \leq m-1$,

$$\begin{aligned} t^{\Delta+kr-1}_{r-1,m} &\geq \frac{1}{(r-1)!^3} \cdot \Big(\ln(m + \tbinom{r-1}{2} + \Delta + kr - 1) - \ln(\tbinom{r-1}{2} + \Delta + kr)\Big)^{r-1}. \\ &= \frac{1}{(r-1)!^3} \cdot \Big(\ln(m + \tbinom{r}{2} + \Delta + kr - r) - \ln(\underbrace{1 + \tbinom{r}{2} + \Delta + kr - r}_{=:f(k)})\Big)^{r-1}, \end{aligned} \tag{9}$$

where we have used Pascal's rule to handle the binomial coefficients. Plugging (9) into (8) and furthermore using the simple inequality $kr + \Delta \leq f(k)$, we thus obtain

$$\begin{aligned} t^{\Delta}_{r,m} &\geq \frac{1}{(r-1)!^3} \sum_{k=1}^{m-1} \frac{1}{rf(k)} t^{\Delta+kr-1}_{r-1,m} \\ &\geq \frac{1}{(r-1)!^3} \int_{x=1}^{m} \frac{1}{rf(x)} \Big(\ln(m + \tbinom{r}{2} + \Delta + rx - r) - \ln f(x)\Big)^{r-1} \,\mathrm{d}x \qquad (10) \\ &\geq \frac{1}{(r-1)!^3} \int_{x=1}^{1+(m-1)/r} \frac{1}{rf(x)} \Big(\ln(m + \tbinom{r}{2} + \Delta) - \ln f(x)\Big)^{r-1} \,\mathrm{d}x \qquad (11) \end{aligned}$$

[1] The term "best-case" here, as well as in Theorem 18, refers to the action of the adversary who chooses the starting position of the augmented process. The intended meaning is for the lower bound in Theorem 18 to hold for any choice of starting position.

$$= \frac{1}{r!^3} \cdot \Big[-\Big(\ln(m + \tbinom{r}{2} + \Delta) - \ln f(x)\Big)^r \Big]_{x=1}^{1+(m-1)/r}$$
$$= \frac{1}{r!^3} \cdot \Big(\ln(m + \tbinom{r}{2} + \Delta) - \ln(1 + \tbinom{r}{2} + \Delta)\Big)^r,$$

which proves the theorem. Note that the integrand in (10) is positive everywhere, so we were justified to restrict the range of the integral in (11), effectively dropping negligible terms.

Theorem 13 now follows from Theorem 18 by setting $\Delta = 0$. ◀

6 Conclusion

Outlook. It remains an open question whether one can obtain good upper bounds for the expected number of steps performed by Random-Edge when the corank is bounded. The only non-trivial result at this point remains the $O(\log^2 n)$ bound for the case $r = 2$, which was settled in [6] and which the author has studied further in a more abstract setting in [13]. It might well be that there is a threshold behavior in the sense that Random-Edge performs well for slowly growing r, and badly for quickly growing r. Finally, the same questions can be asked for other simplex pivoting rules that might be easier to analyze.

Remark on the dependence on r. The leading factor $1/r!^3$ in Theorem 18 is rather small. Due to the results by Friedmann et al. [3], this factor cannot in general be tight. Unfortunately, an improvement seems to be beyond the scope of our method. For the interpretation of Theorem 2 some readers may find it interesting to pick a value of r that depends on the number of facets n. Not every such choice of r leads to a meaningful bound; but it is possible to choose $r = r(n) = \ln^{1/s}\left(1 + 4n/r^4\right)$ with any $s > 3$, which leads to a lower bound of the form $(\ln n)^{\Omega(\ln^{1/s} n)}$ for the directed random walk on a grid polytope with n facets and corank $r(n)$.

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. *Inf. Process. Lett.*, 56(2):79–81, 1995.

2 Stefan Felsner, Bernd Gärtner, and Falk Tschirschnitz. Grid orientations, $(d, d+2)$-polytopes, and arrangements of pseudolines. *Discrete & Computational Geometry*, 34(3):411–437, 2005.

3 Oliver Friedmann, Thomas Dueholm Hansen, and Uri Zwick. Subexponential lower bounds for randomized pivoting rules for the simplex algorithm. In *Proceedings of the 43rd ACM Symposium on Theory of Computing (STOC)*, pages 283–292, 2011.

4 Bernd Gärtner, Martin Henk, and Günter M. Ziegler. Randomized simplex algorithms on klee-minty cubes. *Combinatorica*, 18(3):349–372, 1998.

5 Bernd Gärtner, Walter D. Morris, Jr., and Leo Rüst. Unique sink orientations of grids. *Algorithmica*, 51:200–235, 2008.

6 Bernd Gärtner, József Solymosi, Falk Tschirschnitz, Pavel Valtr, and Emo Welzl. One line and n points. *Random Structures & Algorithms*, 23(4):453–471, 2003 (preliminary version at STOC 2001).

7 Branko Grünbaum. *Convex polytopes*. Springer, New York, 1967/2003.

8 Kathy Williamson Hoke. Completely unimodal numberings of a simple polytope. *Discrete Applied Mathematics*, 20(1):69–81, 1988.

9 Volker Kaibel, Rafael Mechtel, Micha Sharir, and Günter M. Ziegler. The simplex algorithm in dimension three. *SIAM Journal on Computing*, 34(2):475–497, 2005.

10 Gil Kalai. A subexponential randomized simplex algorithm. In *Proc. 24th ACM Symposium on Theory of Computing*, pages 475–482, 1992.

11 Jiří Matoušek and Tibor Szabó. Random edge can be exponential on abstract cubes. *Advances in Mathematics*, 204(1):262–277, 2006.

12 Peter McMullen. On the upper-bound conjecture for convex polytopes. *Journal of Combinatorial Theory, Series B*, 10(3):187–200, 1971.

13 Malte Milatz. Directed random walks on polytopes with few facets. *Electronic Notes in Discrete Mathematics*, 61:869–875, 2017. The European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB '17).

14 Julian Pfeifle and Günter M. Ziegler. On the monotone upper bound problem. *Experimental Mathematics*, 13(1):1–12, 2004.

15 Francisco Santos. A counterexample to the Hirsch Conjecture. *Annals of Mathematics*, 176:383–412, 2012.

16 Falk Tschirschnitz. *LP-related properties of polytopes with few facets.* PhD thesis, ETH Zürich, 2003.