A Markov chain for lattice polytopes

Julien David, Lionel Pournin, Rado Rakotonarivo

February 20, 2018

Abstract. This paper describes an approach to the random sampling of lattice polytopes. The lattice polytopes we are interested in are contained in the hypercube $[0,k]^d$ and we refer to them as lattice (d,k)-polytopes. Our approach consists in using a Markov chain whose space of states is the set of all d-dimensional lattice (d,k)-polytopes and whose transitions add or delete vertices following simple, well-defined rules. We prove that this Markov chain provides a uniform random sampler of lattice (d,k)-polytopes. We also give an upper bound for the mixing time when $d = 2$ and present a number of experimental results on the mixing time for a selection of values for k and d.

1 Introduction

A polytope is the convex hull of a finite set of points in $\mathbb{R}^d$. These objects appear in a wide range of mathematical works, both in theoretical and applied contexts [2], yet their combinatorics is not well understood. A class of polytopes of special interest in combinatorics is that of *lattice (d,k)-polytopes*. These polytopes are contained in the hypercube $[0,k]^d$, where k is a positive integer, and their vertices have integer coordinates. In order to better understand this class of polytopes, we design a random sampler for them using a Markov chain whose space of states is the set of all d-dimensional lattice (d,k)-polytopes.

The transitions in this Markov chain, and the resulting random sampler can be described informally as follows. Given a lattice (d,k)-polytope P with vertex set $\mathcal{V}$, performing a transition will first consist in randomly choosing a lattice point x in $[0,k]^d$, and then to proceed according to the placement of x with respect to P. If x belongs to $\mathcal{V}$, then P will be replaced by the convex hull of $\mathcal{V}\backslash\{x\}$, thus removing x from P. If x does not belong to $\mathcal{V}$ and $\mathcal{V}\cup\{x\}$ is precisely the vertex set of its convex hull, then P will be replaced by the convex hull of $\mathcal{V}\cup\{x\}$, thus inserting x in P. If none of these two cases occurs, then P will not be affected. A formal definition of this Markov chain shall be given in Section 2. In this article we provide both theoretical and experimental results regarding the behaviour of this Markov chain. Our main result is that the resulting random sampler for lattice (d,k)-polytopes is uniform. This is shown in Section 3. Section X concludes the article with a number of experiments on the mixing time of our Markov chain.

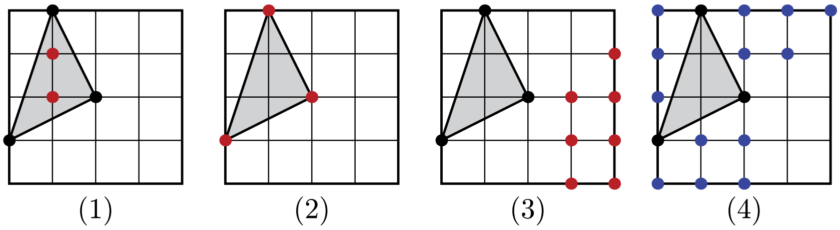

Figure 1: An illustration of the transition rule for a triangle (depicted in grey), depending on the placement of point x, colored red when the chain loops and blue when it does not loop: (1) x belongs to $P\backslash\mathcal{V}$, (2) the convex hull of $\mathcal{V}\backslash\{x\}$ is $(d-1)$-dimensional, (3) $\mathcal{V}\cup\{x\}$ is not the vertex set of its convex hull, and (4) the transition changes P into the convex hull of $\mathcal{V}\cup\{x\}$.

2 The Markov chain

We will consider a Markov chain whose space of states Ω is the set of all d-dimensional lattice (d,k)-polytopes, for fixed d and k. A priori, the definition of lattice (d,k)-polytopes does not requires them to be d-dimensional. Some effort will be needed in order to enforce the requirement that all the states of our Markov chain are polytopes of dimension exactly d. The transition rules of our Markov chain are defined as local operations on lattice (d,k)-polytopes. These rules consist in adding or removing a given vertex to such a polytope. Consider a d-dimensional lattice (d,k)-polytope P and denote by $\mathcal{V}$ its vertex set. Consider a lattice point x in $[0,k]^d$ which, we assume has been uniformly drawn from all possible lattice points in $[0,k]^d$. The transition from P that corresponds to the chosen lattice point x goes as follows.

- If x is contained in P but is not a vertex of it (i.e. $x\in P\backslash\mathcal{V}$) then the chain will loop on P. In other words, the current state is unaffected.
- If x is a vertex of P (i.e. $x\in\mathcal{V}$), then
  - If the convex hull Q of $\mathcal{V}\backslash\{x\}$ is d-dimensional, the transition goes from P to Q. Note that if Q were $(d-1)$-dimensional, then P would be a pyramid (with apex x) over Q. In this case, the transition from P to Q would be impossible because it would exit Ω.
  - Otherwise, we loop on P.
- If x does not belong to P, then
  - If $\mathcal{V}\cup\{x\}$ is precisely the vertex set of its convex hull, then the transition goes from P to the convex hull of $\mathcal{V}\cup\{x\}$.
  - Otherwise we loop on P.

Figure 1 illustrates transition rule in the case of a lattice triangle P contained in the square $[0,4]^2$, depending on the placement of point x. Note that in this particular case, there is no transition that deletes a vertex of P.

3 Properties of the Markov chain

The purpose of this section

The purpose of this paper is to build the random sampler of (d,k)-polytopes, in order to achieve our goal, we need to ensure that the stationnary distribution on Ω is the uniform. An important result on Markov chains shows that an irreducible, aperiodic and symetric Markov chain admits as stationnary distribution the uniform [1].

This section aim to verify these properties on (X_t), and give our first results on (X_t). Only verifying the irreducibility has been slightly tricky, the other ones can be directly proved. In order to do so, the notion of symetric difference on our states will be introduced. Thereby, given x and $y \in \Omega$, we recall that their *symetric difference*, $x \bigtriangleup y$, is defined as:

$$x \bigtriangleup y = x \cup y \setminus x \cap y \tag{1}$$

Recall that the irreducibility mean that all the states of Ω can be reached from any other state, thus the irreducibility guarantees that the graph with vertex set Ω is connected. In fact move from x to y in our chain consists in finding finite number of transitions between x and y by one. Observe that the simplest way to move from x to y consists in adding directly to x the vertices of y then remove x vertices. This means that at each step one reduces the symetric difference between x and y. Hence, one can verify that the cardinality of $x \bigtriangleup y$ is a lower bound for the distance, noted $\delta(x,y)$, between x and y. One has:

$$\delta(x,y) \geq |x \bigtriangleup y| \tag{2}$$

However, we might face cases which not allow this fast transition, thus we have to find out other transition state instead of those *trivial* cases. Let us take into account the difficult cases and introduce lemmas to build our proof to the irreducibility. Not being able to add a point $w \in y \setminus x$ to x means that for all $w \in y \setminus x$ the following cases occur: w is an interior point to x, w belongs to the affine hull of the edge of x, v belongs to the cone described by one vertex of x and the two edges which contains this vertex. To overcome this problem, we claim that we can always find a point v in $[0,k]^d$ we can add to x, have a new state from wich we are building the path to y, then ensure the irreducibility.

3.1 Existence of v

Consider a d-dimensional lattice (d,k)-polytope $\mathcal{S}$ and assume that $\mathcal{S}$ is a simplex. Let $\mathcal{V}$ be the set of vertices of $\mathcal{S}$. For any $u \in \mathcal{V}$, denote by H_u the affine hull of the facet of $\mathcal{S}$ opposite u, and by H_u^- the closed half-space of $\mathbb{R}^d$ limited by H_u that does not contain u. For any $v \in \mathcal{V}$, the set

$$C_v = \bigcap_{u \in \mathcal{V} \setminus \{v\}} H_u^- \tag{3}$$

is a d-dimensional cone pointed at v. This cone is exactly the set of the points $x \in \mathbb{R}^d$ such that the convex hull of $\mathcal{S} \cup \{x\}$ does not admit v as a vertex. By this remark, we have the following lemma.

Lemma 3.1.1. *Let x be a lattice point in $[0,k]^d$. The convex hull of $\mathcal{S}\cup\{x\}$ admits, as vertices, x and all the vertices of $\mathcal{S}$ if and only if x does not belong to $\mathcal{S}$ and, for all $v\in\mathcal{V}$, x does not belong to C_v.*

Call $\gamma=\min\{x_1 : x\in S\}$. Consider the following hyperplane of $\mathbb{R}^d$:

$$X=\{x\in\mathbb{R}^d : x_1=\gamma\}$$

Denote by X^- the open half-space of $\mathbb{R}^d$ bounded by X and that does not contain $\mathcal{S}$. By construction, the intersection $\mathcal{S}\cap X$ is a non-empty, proper face of $\mathcal{S}$. This face will be denoted by F in the following. Since $\mathcal{S}$ is a simplex, it admits another, non-empty face $F^\star$ whose vertices are exactly the vertices of $\mathcal{S}$ that do not belong to F.

By construction,

$$\dim(F)+\dim(F^\star)=d-1.$$

In particular, there exists a vector c that is orthogonal to both F and $F^\star$. Consider the hyperplane Y of $\mathbb{R}^d$ that admits c as a normal vector and such that $F^\star\subset Y$. The intersection $\mathcal{S}\cap Y$ is precisely $F^\star$. Denote by Y^- the closed half-space of $\mathbb{R}^d$ bounded by Y that does not contain F. It will be assumed that c has norm 1 and that it points towards Y^-. Let

$$\delta=\min\{c\cdot x : x\in X\cap[0,k]^d\}. \tag{4}$$

Further denote $G=\{x\in X\cap[0,k]^d : c\cdot x=\delta\}$. In the statement of the following lemma, $\operatorname{aff}(F)$ denotes the affine hull of F.

Lemma 3.1.2. *If v is a vertex of $\mathcal{S}$, then*

$$C_v\subset\operatorname{aff}(F)\cup X^-\cup Y^-.$$

Proof. First observe that, if s is a face of $\mathcal{S}$ and H is a hyperplane of $\mathbb{R}^d$ that intersects $\mathcal{S}$ exactly along s, then

$$\bigcap_{u\in\mathcal{V}\setminus s} H_u^-\subset H^-,$$

where H^- is the closed half space of $\mathbb{R}^d$ bounded by H and disjoint from the interior of $\mathcal{S}$. Further observe that the intersection of H with

$$\bigcap_{u\in\mathcal{V}\setminus s} H_u^-$$

is precisely the affine hull of $\mathcal{S}$. As a direct consequence, taking in turn $s=F^\star$ and $s=F$, one obtains that, if v is a vertex of F, then $C_v\subset\operatorname{aff}(F)\cup X^-$ and if v is a vertex of $F^\star$, then $C_v\subset Y^-$. The result therefore holds because any vertex of $\mathcal{S}$ is either a vertex of F or a vertex of $F^\star$. ■

As an immediate consequence of this and Lemma 3.1.1, for any lattice point $x\in[0,k]^d$ that does not belong to X^-, to Y^-, or to the affine hull of F, the convex hull of $\mathcal{S}\cup\{x\}$ admits, as vertices, x and all the vertices of $\mathcal{S}$.

Recall that c is orthogonal to F and $F^\star$. As a consequence, the map $x\mapsto c\cdot x$ is constant within F and within $F^\star$. Call ε the value of $c\cdot x$ when $x\in F$ and $\varepsilon^\star$ the value of $c\cdot x$ when $x\in F^\star$. Since F and Y^- are disjoint, $\varepsilon<\varepsilon^\star$. Moreover, by (4), $\delta\leq\varepsilon$. Observe that if the latter inequality is strict, then G is disjoint from both $\operatorname{aff}(F)$ and Y^-. By definition, it is also disjoint from X^- and the following lemma is then obtained as a consequence of Lemmas 3.1.1 and 3.1.2.

Lemma 3.1.3. *If $\delta < \varepsilon$, then for any lattice point $x \in G$, the convex hull of $\mathcal{S} \cup \{x\}$ admits, as vertices, x and all the vertices of $\mathcal{S}$.*

If, on the contrary, δ and ε coincide, then $F \subset G$. This situation is familiar: we are looking at a lattice simplex F contained in a (possibly degenerate) lattice cube G. If the dimension of G is greater than the dimension of F, then the following lemma provides the desired result.

Lemma 3.1.4. *If k and d are positive and if P is a lattice (d, k)-polytope of dimension less than d then there exists a lattice point x that belongs to $[0, k]^d$ but that does not belong to the affine hull of P.*

Proof. If P is a lattice (d, k)-polytope of dimension less than d, then the intersection I of its affine hull with $[0, k]^d$ cannot contain more than $(k+1)^{d-1}$ lattice points. Indeed, one can always project I orthogonally on a facet of $[0, k]^d$ in such a way that the dimension of the projection is not less than that of I. Such a projection induces an injection from the lattice points in I into the lattice points in the facet on which the projection is made.

Now observe that $[0, k]^d$ contains $(k + 1)^d$ lattice points. Since k is positive, $(k + 1)^{d-1} < (k + 1)^d$ and the lemma is proven. ■

Hence, it remains to solve the case when F is a subset of G and both have the same dimension. If this dimension is at least 2, then the strategy is to argue by induction on d. The base case of the induction is given by the following lemma.

Lemma 3.1.5. *If $d = 2$ then there exists a lattice point $x \in [0, k]^2$ such that the convex hull of $\mathcal{S} \cup \{x\}$ admits, as vertices, x and all the vertices of $\mathcal{S}$.*

Proof. Probably just a careful, hopefully short disjunction. Rado, Julien? ■

The case when F is a subset of G and their common dimension is either 0 or 1 has to be treated separately.

Lemma 3.1.6. *Assume that d is greater than 2. If G has dimension at most 1 and admits F as a subset, then there exists at least one lattice point x in $[0, k]^d$ that does not belong to $\operatorname{aff}(F) \cup X^- \cup Y^-$.*

Proof. Assume that the dimension of G is 0 or 1 and that F is a subset of G. Because of the latter assumption, $\delta = \varepsilon$. Further assume that every lattice point in $X \cap [0, k]^d$ belongs to either $\operatorname{aff}(F)$ or Y^-. Since Y^- is the set of points $x \in \mathbb{R}^d$ such that $c{\cdot}x \geq \varepsilon^\star$, this assumption yields that any lattice point x in $X \cap [0, k]^d$ such that $c{\cdot}x < \varepsilon^\star$ satisfies $c{\cdot}x = \delta$.

In particular, the only lattice points in $X \cap [0, k]^d$ that may belong to Y have distance exactly 1 to some lattice point in G.

Now consider the set N of the points x in $[0, k]^d$ whose orthogonal projection on X belongs to G and such that $x_i = \gamma + 1$. By construction, $\gamma < k$ and therefore, N is non-empty. More precisely, N is made up of a single point if G has dimension 0, and N is a line segment parallel to G if G has dimension 1. In particular, the map $x \mapsto c{\cdot}x$ is constant within N. Call this constant η. If $\eta \geq \varepsilon^\star$, then the only lattice points x in $[0, k]^d$ such that $x_i \geq \gamma$ that may belong to Y are the ones whose distance to some lattice point in G is exactly 1. Among these candidates, the only ones that can also be vertices of $\mathcal{S}$ are the lattice point in N. Indeed, all the other candidates belong to X. This is impossible because $\mathcal{S}$ would then be contained in the convex hull of $G \cup N$ whose dimension is either 1 or 2 depending on the dimension of G, whereas the dimension of $\mathcal{S}$ is at least 3.

It follows from this contradiction that $\eta < \varepsilon^\star$. In other words, N is disjoint from Y^-. By construction, N is also disjoint from X^- and from the affine hull of F. As N contains lattice points, this proves the lemma. ■

Here now comes the following proposition we want to claim.

Proposition 3.1.1. *There exists a lattice point x in the cube $[0,k]^d$ such that the convex hull of $\mathcal{S} \cup \{x\}$ admits, as vertices, x and all the vertices of $\mathcal{S}$.*

Proof. We need to write the induction carefully. Rado, Julien? ■

3.2 Irreducibility of (X_t)

Lemma 3.2.1. *For any simplex $x \in \Omega$, for any $y \in \Omega$. If we cannot reduce $x \triangle y$ by adding a vertex in $y \setminus x$ the there exists a simplex $z \in \Omega$, a transition state between x and y, which verifies:*

- $|x \triangle y| = |z \triangle y|$
- $\delta(x,z) \leq 2$

such that we can always add a vertex in $y \setminus z$, from z to y.

Proof. Let be x a simplex y be in Ω. Let be $v \in [0,k]^d$. The idea of the proof lies on the fact that, z will be found by adding a point $v \in \Omega$ and which is not an element of $x \triangle y$, then remove a point of $x \setminus y$. In this way we found the simplex z at most in 2 steps. ■

Lemma 3.2.2. *For all x and $y \in \Omega$, there exists $z \in \Omega$ which satisfies $|x \triangle y| > |z \triangle y|$ such that $\delta(x,z) \leq 3$.*

Proof. Let x and y be states of Ω, such that $P(x,y) = 0$. Let z be a transition state between x and y, and such that z is closer to y than y. We have several distinct cases:

1. x is not a simplex.
   (a) $x \subset y$: add $v \in y \setminus x$ and $z = x \cup \{v\}$, then $\delta(x,z) = 1$
   (b) $x \not\subset y$: remove $v \in x \setminus y$ and $z = x - \{v\}$, then $\delta(x,z) = 1$
2. x is a simplex.
   (a) If we can add $v \in y \setminus x$ then add it, thus $z = x \cup \{v\}$ and $\delta(x,z) = 1$
   (b) Else:
      i. Add a point u which is not in $x \triangle y$
      ii. Remove a point from $x \setminus y$
      iii. Add a point from $y \setminus x$
      In this last case, one finds z such that $\delta(x,z) = 3$

Since we had proved we can always add a point which is not contained in $x \triangle y$ by lemma 3.2.1. It means that wherever in which case we are, one can always find a z which reduces the symetric difference from z to y by one, such that $\delta(x, z) \leq 3$.

■

Corollary 3.2.1. *For all x and $y \in \Omega$ one has:*

$$\delta(x, y) \leq |x| + |y| + 4(d + 1) \tag{5}$$

Proof. This is an immediate consequence of the lemma 3.2.2. Let x and y be in Ω. Let us now consider two simplices $x^\star$ and $y^\star$ such that $\delta(x, x^\star) = |x| - (d+1)$, and $\delta(y, y^\star) = |y| - (d + 1)$. Thus

$$\delta(x, y) \leq \delta(x, x^\star) + \delta(x^\star, y^\star) + \delta(y, y^\star)$$

Since $x^\star$ is a simplex, the walk needs at most $3(|x^\star| + |y^\star|) = 3 \times 2(d + 1)$ steps to reach $y^\star$ from $x^\star$. Hence

$$\delta(x, y) \leq |x| - (d + 1) + |y| - (d + 1) + 6(d + 1) = |x| + |y| + 4(d + 1)$$

■

Observe that corollary 3.2.1 gives an idea on the upper bound of the diameter of (X_t). We have now settled all we need to prove our main results on the properties of (X_t).

Theorem 1. *Define the diameter $\mathcal{D}$ of a graph with vertex set Ω to be the maximum distance between two vertices. For (X_t), as defined above, and given k and d, one has:*

$$\mathcal{D}_{X_t} \leq 2ck^{3/4} + 4(d + 1) \quad \textit{where } c > 0 \tag{6}$$

Proof. To be done. ■

Theorem 2. *The Markov chain (X_t) is irreductible, aperiodic and has the uniform as stationnary distribution.*

Proof. Three propreties have to be verified, thus this proof will be given in three steps. Let x and y be in Ω.

i *Irreducibility* Irreducibility is a direct consequence of the corollary 3.2.1. Let us remind that to prove the irreducibility, one needs to find a r_0 such that, for all x and $y \in \Omega$, when $r \geq r_0$ then $P^r(x, y) > 0$. Thus let us take $r_0 = |x| + |y| + 4(d + 1)$.

ii *Symetry* Our transition rules consists in either add or remove a single vertex for two distinct states. Observe that if there is no one step transition from x to y means $P(x, y) = 0$. By the trasition rules, necessarily, if $P(x, y) = 0$ then $P(y, x) = 0$. Next, let us prove that if $P(x, y) > 0$ and $P(y, x) > 0$ then we also have $P(x, y) = P(y, x)$. $P(x, y) > 0$ means that we have a one step transition from x to y. Only two cases may occur. For $v \in [0, k^d]$ either $y = x - \{v\}$, or $y = x \cup \{v\}$. Considering the first case, the probability to draw v and add it to x is $P(x, y) = \frac{1}{(k+1)^d}$. Similarily, to move from y to x, we draw the same v and remove it from y to get x with probability $P(y, x) = \frac{1}{(k+1)^d} = P(x, y)$. We prove the remaining case in an analog reasoning.

iii *Aperiodicity* To prove that (X_t) is aperiodic, one needs to show that each state in Ω has as period 1. Since (X_t) is irreductible, property **??** tells us that all the states of Ω has the same period. Thus, for all $x, y \in \Omega$, $\gcd(\mathcal{T}(x)) = \gcd(\mathcal{T}(y))$. One needs to find a state x such that $\text{tgcd}(\mathcal{T}(x)) = 1$. Let us take x has a simplex and a point $v \in [0, k]^d$ then consider the cases where we have a loop on x: either v is an interior point to x, or v is drawn outside of x but $|Conv(x \cup \{v\})| \neq |x|+1$. One has: $P(x, x) = \mathbf{P}\{v \text{ interior point to } x\} + \mathbf{P}\{v \in x\} + \mathbf{P}\{|x \cup \{v\}| \neq |x| + 1\}$. Note that $\mathbf{P}\{v \in x\} = \frac{1}{(k+1)^d}$, but since $|x| = d + 1$, we have a loop on x with probability $P(x, x) \geq \frac{1}{(k+1)^d} > 0$. In another words, with positive probability the walk can get back to x from x in one step. Hence, $\mathcal{T}(x) = \{1, \ldots\}$. We conclude that $\gcd(\mathcal{T}(x)) = 1$.

■

4 Random sampler

Consider (X_t) as defined previously. Sampling random (d, k)-polytopes consists in a random walk on Ω with our transition rules until one reaches the stationnary distribution. The amount of time needed to reach such a distribution, that is sample a uniformally random (d, k)-polytope, is the mixing time on (X_t).

Sampling a random (d, k)-polytope with this model is given the following algorithm.

Algorithm 1: $\Gamma(d, k)$: Sample a (d, k)-polytope

Input: the dimension d, the side k of the hypercube
Output: random (d, k)-polytope drawn in $[0, k]^d$

```
1 arbitrary initial state x_0
2 while the stationnary distribution is not reached do
3     draw a random point v in [0,k]^d
4     if v ∈ x_0 et |x_0| > d+1 then
5         x_0 = x_0 − {v}
6     else if |Conv(x_0 ∪ {v})| == |x_0| + 1 then
7         x_0 = x_0 ∪ {v}
8 return x_0
```

Theorem 3. *The random sampler $\Gamma(d, k)$ described by algorithm 1 is a uniform random sampler of (d, k)-polytopes over Ω.*

5 Results on mixing time

References

[1] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times.* American Mathematical Soc., 2009.

[2] Günter M Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 1995.