Pointed *versus* Singular Boltzmann Samplers*

Olivier Bodini†, Antoine Genitrini‡§ and Nicolas Rolin†.
{Olivier.Bodini;Nicolas.Rolin}@lipn.univ-paris13.fr and Antoine.Genitrini@lip6.fr.

May 21, 2014

For the last twenty years, giant graphs and huge systems have played a central role in computer science because of the technological improvements. In order to study such structures, an enthusiasm for random generation has emerged. A new technique has been introduced ten years ago: the Boltzmann sampling. It has been presented by Duchon *et al*, and is based on automatic interpretation in terms of samplers of the specification of the combinatorial objects under study.

One of the core problem in Boltzmann sampling lies in the distribution of the object sizes, and the choice of some parameters in order to get the more appropriate size distribution. From this choice depends the efficiency of the sampling. Moreover some additional ideas allows to improve the efficiency, one of them is based on some anticipated rejections, the other one on the combinatorial differentiation of the specification. Anticipated rejection consists during the recursive building of a random object to kill the process as soon as we are sure to exceed the maximum target size, rather than waiting until the natural end of the process. In the original paper, the two different methods are proposed. But, while both approaches have been presented in that paper, and used on the same kind of structures, the methods are not compared.

We propose in this paper a detailed comparison of both approaches, in order to understand precisely which method is the more efficient.

1 Introduction

For the last two decades, giant graphs and huge systems play a central role in computer science due to the technological improvements. In order to study such structures, an enthusiasm for random generation has emerged. In fact, such an approach allows, for example, to do testing: uniform generation of huge graphs enables to detect errors [CD09], particularly to detect unexpected overflows that wouldn't have been given away on smaller instances. Another application is to create instances to understand how the structures work: let us mention the contexts of group theory [BNW08], formal languages [BN07] and statistical physics [BFP10].

The first algorithms that uniformly sample combinatorial structures are *ad hoc* methods which are dedicated to a specific class of objects. Let us, for example, mention the algorithm of Rémy [Rém85], recently improved in [BBJ14] that allows to sample uniformly binary trees. Another usual way to sample uniformly at random combinatorial structures is based on the recursive method developed in the first part of the book [NW78] and then revisited in the context of Analytic Combinatorics [FZC94]. This approach relies on a systematic process to generate combinatorial objects.

*This research was partially supported by the A.N.R. project *MAGNUM*, ANR 2010-BLAN-0204.

†Laboratoire d'Informatique de Paris-Nord, CNRS UMR 7030 - Institut Galilée - Université Paris-Nord, 99, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.

‡Sorbonne Universités, UPMC Univ Paris 06, UMR 7606, LIP6, F-75005, Paris, France.

§CNRS, UMR 7606, LIP6, F-75005, Paris, France.

Most of the time it is a bit less efficient that *ad hoc* methods, however it gets the nice property to adapt to many families of objects that are usually qualified of decomposable (or specifiable) objects.

Since 2004, a new technique has been introduced: the Boltzmann sampling. It has been presented by Duchon *et al*, in [DFLS04], and is based on the decomposition of the combinatorial objects under study, through their specification. The approach has, then, been widely explored in order to fit to much more constructions, see the papers [FFP07, Duc11, BRS12, BP10] to have an idea of some of the new developments. Boltzmann sampling plays a central rôle in the context of random generation, because of its simplicity, its genericity and its efficiency. All these good properties are however combined with a constraint: in order to be efficient, the generator builds objects with an approximate size instead of the size fixed by the user. In many cases, this constraint is not a problem at all: when one is interested, for example, to test a huge system, it does no matter if the size of its simulation is around one million instead of being exactly one million.

The Boltzmann samplers need a parameter in order to generate objects. The tuning of this parameter allows for example to build objects whose expected size could be the fixed size, one wants to reach. Then, once the parameter is chosen, a tricky point is the evaluation of generating series in specific values. Some studies are presenting improvements, based on the quadratic iterative Newton method [PSS12], or a curve approach of the problem [BLR13] in order to overcome these difficulties.

One of the core problem in Boltzmann sampling lies in the distribution of the object sizes, and the choice of the parameter in order to get the best size distribution. An interesting distribution would have its mass concentrated around its mean and the selected parameter would yield a good probability that a randomly generated object falls in a size range near the mean. But some combinatorial structures do not exhibit such nice distributions, and have rather their distribution mass around extreme values, like distribution with "big tails". Even a smart choice of the parameter does not give an interesting complexity.

A first step to deal with these problems is based on some anticipated rejection. It consists during the recursive building of a random object in killing the process as soon as we are sure (by the size of the under construction structure) to exceed the maximum target size, rather than waiting until the natural end of the process. In [DFLS04], two different methods are proposed to be applied to big tails distributions. (1) A method based on some pointing: the size distribution is modified in order to obtain some more interesting distribution. (2) A method based on singular sampling: rejection is used and the parameter is taken to be the dominant singularity of the generating series, consequently the choice of the parameter and the heavy evaluation of the series is avoided. While both approaches have been presented, in [DFLS04], on the same kind of structures; only the complexities for the specific case of the singular exponent of simple family of trees have been computed, but without comparison. It isn't clear in the litterature which method is the more efficient. We propose in this paper a detailed comparison of both approaches in more general cases (not only for the singular exponent of simple family of trees).

The paper is organized as follows. Section 2 gives the contextual definitions and notions about Boltzmann sampling. Sections 3 and 4 analyse the time complexity of the different methods of sampling. Finally, Section 5 is devoted to the comparison, in terms of time complexity, of the two methods based on pointing or singular sampling, first in the general structures and then in the case of aperiodic and strongly connected grammars (singular exponent $-1/2$).

2 Context of Boltzmann samplers

In this section, we recall the classical definitions in the context of combinatorial structures and Boltzmann samplers. In all the paper we will deal with unlabelled combinatorial classes (associated to ordinary generating functions). However, the adaptation to the labelled structures is straightforward.

2.1 Boltzmann distributions

A *combinatorial class* $\mathcal{C}$ is a set of objects, with a size function, denoted by $|\cdot| : \mathcal{C} \to \mathbb{N}$ and such that for every integer n, the subset $\mathcal{C}_n$ of objects of size n, is finite with cardinality C_n. We define the *ordinary generating function* of a combinatorial class $\mathcal{C}$ to be:

$$C(z) = \sum_{n\geq 0} C_n z^n = \sum_{\gamma\in\mathcal{C}} z^{|\gamma|}.$$

Let ρ, the *dominant singularity* of the generating function $C(z)$.

Definition 2.1. *Let $\mathcal{C}$ be an unlabelled combinatorial class, whose ordinary generating function is $C(z)$. Its associated Boltzmann model, depending on a real parameter $0 < x < \rho$ or $x = \rho$ in some cases, is the distribution over the objects of $\mathcal{C}$, such that the probability of an object γ is*

$$\mathbb{P}_x(\gamma) = \frac{x^{|\gamma|}}{C(x)}, \qquad \text{where } C(x) \text{ is the generating function of } \mathcal{C} \text{ evaluated in } x.$$

We note that the distribution is uniform among all objects of the same size. Hence, the probability of getting an object of size n, denoted by $\mathbb{P}_x(N = n)$ and the expected size, denoted by $\mathbb{E}_x(N)$, of the object, for the parameter x, satisfy:

$$\mathbb{P}_x(N = n) = \frac{C_n x^n}{C(x)}, \qquad \text{and} \qquad \mathbb{E}_x(N) = \frac{x\, C'(x)}{C(x)}.$$

2.2 Boltzmann samplers

A *Boltzmann sampler* is a generator of objects according to their Boltzmann distribution. Since the size of a generated object is random, in order to sample an object of a given size n, the samplings are repeated until getting an object of the right size. To obtain a generation [DFLS04] that is more efficient, usually, objects with an approximation on the size are accepted. For example, an object whose size belongs to the range $n(1-\varepsilon)$ and $n(1+\varepsilon)$ is accepted, where ε is the tolerance on the size.

We will restrict the study to combinatorial classes whose generating functions are $\Delta - singular$. Remark that $\Delta - singular$ functions are associated to combinatorial decomposable objects, as argued in [FZC94].

Definition 2.2. *A function $C(z)$ analytic at 0 and with a finite radius of analyticity $\rho > 0$ is said* Δ-singular *if it satisfies the two following conditions.*

(i) The function admits ρ as its only singularity on $|z| = \rho$ and it is continuable in a domain

$$\Delta(r, \theta) = \{z \mid z \neq \rho, |z| < r, arg(z - \rho) \notin (-\theta, \theta)\},$$

for some $r > \rho$ and some θ satisfying $0 < \theta < \frac{\pi}{2}$. The set $\Delta(r, \theta)$ is called a Δ-domain.

(ii) For z tending to ρ in the Δ-domain, $C(z)$ satisfies a singular expansion of the form

$$C(z) =_{z\to\rho} P(z) + c_0(1 - z/\rho)^{-\alpha}) + o((1 - z/\rho)^{-\alpha}), \qquad \alpha \in \mathbb{R}\backslash\{0, -1, -2, ...\},$$

where $P(z)$ is a polynomial; c_0 a constant and $-\alpha$ is called the singular exponent *of $C(z)$.*

Using the Flajolet-Odlysko transfer theorems, (cf. [FS09, chapter VI] for details), the coefficients of Δ-singular functions satisfy the following asymptotic behaviour:

$$C_n := [z^n]C(z) \sim \frac{c_0}{\Gamma(\alpha)} \rho^{-n} n^{\alpha-1}.$$

Note that for $-\alpha > 0$, the size distributions of objects have a heavy tail: such distributions are usually called *peaked distributions.* While the distributions induced by $-\alpha < 0$ are called *flat* *distributions.*

In order to generate objects according to their Boltzmann distribution, Duchon *et al* [DFLS04] have presented three samplers. The first one, that we will call *the basic Boltzmann sampler*, is based on the best parameter x_n as possible, in order to maximize the chances to get an object of the appropriate size. The second sampler, called *pointed Boltzmann sampler* deals with flat size distributions of objects. In order to turn a peak distribution (induced by trees for example) to a flat distribution, we can mark some atoms of the structures, then generate such marked structures and finally avoid the marked nodes. Obviously, this method is still uniform among objects of the same size, even if the global distribution is now different from the original Boltzmann one. Finally, the third generator, called *singular Boltzmann sampler*, is based on peaked size distributions of structures. In this case, we can take the dominant singularity as parameter.

2.3 Complexity of Boltzmann samplers

For a Boltzmann sampler, the complexity is usually measured by the cumulated size of the objects that are generated and then left because their sizes are not in the good range. This measure is relevant, because in Boltzmann sampling, the time, the space and the number of random bits that are used is growing linearly with the size of the sampled objects (cf. [DFLS04]).

Considering such a complexity measure, an obvious improvement of the sampler, consists of not building the objects larger that the maximum size fixed. So by rejecting an object as soon as one of its part reaches the maximum size. This procedure is called *anticipated rejection*. Let us first present a result proved in [DFLS04], but that is central for our study. It gives the complexity of the approximate Boltzmann sampling with anticipated rejection.

Lemma 2.1. *[DFLS04] Let $C(z)$ be the generating function of a class $\mathcal{C}$, and let $C^{<n_1}$,$C^{>n_2}$ and $C^{[n_1,n_2]}$ be the generating functions for the subclasses of objects with size, respectively, strictly smaller than n_1, strictly greater than n_2, and between n_1 and n_2. The cumulative size, T_n, of the generated and rejected objects with a tolerance window at $[n_1, n_2]$ and with parameter x satisfies:*

$$\mathbb{E}(T_n) = \frac{xC'^{<n_1}(x) + n_2C^{>n_2}(x)}{C^{[n_1,n_2]}(x)}.$$

Without anticipated rejection, we get:

$$\mathbb{E}(T_n) = \frac{xC'^{<n_1}(x) + xC'^{>n_2}(x)}{C^{[n_1,n_2]}(x)}.$$

Proof. The following proof is given in [DFLS04].
The probability generating function of the approximate Boltzmann sampler with rejection targeted at $[n_1, n_2]$ is

$$F(u, x) = \sum_k \mathbb{P}(T_n = k)u^k.$$

From the decomposition of a call to the sampler into a sequence of unsuccessful trials, (each one contributing to T_n) and then followed by a final successful trial (not contributing to T_n), we get:

$$F(u, x) = \left(1 - \frac{1}{C(x)}(C^{<n_1}(ux) + C^{>n_2}(x)u^{n_2})\right)^{-1} \frac{C^{[n_1,n_2]}(x)}{C(x)}.$$

Then the expectation of the cost is given by $\mathbb{E}(T_n) = \frac{\partial d}{\partial du}F(u, x)_{|u=1}$.
By observing that $C(x) - C^{<n_1}(x) - C^{>n_2}(x) = C^{[n_1,n_2]}(x)$, the result follows immediately. □

This lemma allows to compute the expected cost of an approximate Boltzmann sampler with anticipated rejection as long as the quantities $xC'^{<n_1}(x)$, $n_2C^{>n_2}(x)$ and $C^{[n_1,n_2]}(x)$ can be calculated.

We end this section by recalling some notations associated to each sampler, that have been introduced in [DFLS04]. Let $\mu C(x,n,\varepsilon)$ represents the Boltzmann sampler for the combinatorial class $\mathcal{C}$, without anticipated rejection, with parameter x, size n and size-tolerance ε. And let $\nu C(x,n,\varepsilon)$ be the analogous Boltzmann sampler with anticipated rejection.

We turn now to the computations of the sampler complexities. In order to classify our study, we partition the combinatorial structures according to the value of their singular exponent.

3 Objects presenting a flat size distribution

In this section, we are considering combinatorial structures, whose singular exponent satisfies $-\alpha < 0$. Such objects present a size distribution that is flat. For example, tree structures with a marked node belong to this class of objects.

To sample an object of size n, the canonical way of choosing an appropriate parameter x_n is such that the expected size of the generated objects equals or is close to n. Thus let x_n be the solution of $\mathbb{E}_{x_n}(N) = x_n$. In the case when $-\alpha < 0$, we get $x_n \sim \rho(1-\frac{\alpha}{n})$, hence we choose x_n to be equal to $\rho(1-\frac{\alpha}{n})$.

3.1 Basic Boltzmann sampler without anticipated rejection

Theorem 3.1. *Let $\mathcal{C}$ be a combinatorial class whose generating function is Δ-singular with an exponent $-\alpha < 0$. Then the cumulated size T_n of the objects generated and rejected by an approximate Boltzmann sampler without anticipated rejection, $\mu C(x_n,n,\varepsilon)$, satisfies, when n tends to infinity:*

$$\mathbb{E}(T_n) \sim n\kappa(\varepsilon,\alpha), \qquad \text{where } \kappa(\varepsilon,\alpha) = \frac{\int_{w=0}^{1-\varepsilon} w^{\alpha} e^{-\alpha w} dw + \int_{w=1+\varepsilon}^{\infty} w^{\alpha} e^{-\alpha w} dw}{\int_{w=1-\varepsilon}^{1+\varepsilon} w^{\alpha-1} e^{-\alpha w} dw}.$$

Proof. In order to use the Lemma 2.1, we need estimations for $x_n C'^{<n(1-\varepsilon)}(x_n)$, $x_n C'^{>n(1+\varepsilon)}(x_n)$ and $C^{[n(1-\varepsilon),n(1+\varepsilon)]}(x_n)$. We use an Euler-MacLaurin summation on the coefficients of C to obtain those estimations:

$$x_n C'^{<n(1-\varepsilon)}(x_n) \sim \frac{c_0 n^{\alpha+1}}{\Gamma(\alpha)} \int_{w=0}^{1-\varepsilon} w^{\alpha} e^{-\alpha w} dw, \qquad x_n C'^{>n(1+\varepsilon)}(x_n) \sim \frac{c_0 n^{\alpha+1}}{\Gamma(\alpha)} \int_{w=1+\varepsilon}^{\infty} w^{\alpha} e^{-\alpha w} dw,$$

and

$$C^{[n(1-\varepsilon),n(1+\varepsilon)]}(x_n) \sim \frac{c_0 n^{\alpha}}{\Gamma(\alpha)} \int_{w=1-\varepsilon}^{1+\varepsilon} w^{\alpha-1} e^{-\alpha w} dw.$$

□

Let us now turn to the sampler with anticipated rejection, in order to be able to explicit if the approach is much more efficient.

3.2 Basic Boltzmann sampler with anticipated rejection

In that context, during the generation of a random object, we compute the size of its components under construction. Once the size is larger than the maximum allowed, reject the object under construction, and start a new one.

Theorem 3.2. *Let $\mathcal{C}$ be a combinatorial class whose generating function is Δ-singular with an exponent $-\alpha < 0$. Then the cumulated size T_n of the objects generated and rejected by an approximate Boltzmann sampler with anticipated rejection, $\nu C(x_n, n, \varepsilon)$, satisfies, when n tends to infinity:*

$$\mathbb{E}(T_n) \sim n\kappa_c(\varepsilon, \alpha), \qquad \text{where } \kappa_c(\varepsilon, \alpha) = \frac{\displaystyle\int_{w=0}^{1-\varepsilon} w^{\alpha} e^{-\alpha w} dw + \int_{w=1+\varepsilon}^{\infty} (1+\varepsilon) \cdot w^{\alpha-1} e^{-\alpha w} dw}{\displaystyle\int_{w=1-\varepsilon}^{1+\varepsilon} w^{\alpha-1} e^{-\alpha w} dw}.$$

The proof deeply looks like the proof of Theorem 3.1.

4 Objects presenting a peaked size distribution

In this section, we are considering combinatorial structures, whose singular exponent satisfies $-\alpha > 0$. Such objects present a size distribution that is peaked. Tree structures are examples of this kind of objects.

4.1 Pointed Boltzmann sampler

A pointed object is a structure with a node that is marked. Thus, a pointing Boltzmann sampler generates a pointed structure, but returns the structure without the mark. This operation breaks the symmetry during the generation and turns effectively peaked distributions into flat distributions. Of course, the generation is still uniform among objects of the same size. Given a combinatorial class $\mathcal{C}$, we define the pointed class (associated to any kind of traversal) is defined as:

$$\mathcal{C}^{\bullet} \sim \{(\gamma, i) \mid \gamma \in \mathcal{C}, i \in \{1, \ldots, |\gamma|\}.$$

We deduce $|\mathcal{C}_n^{\bullet}| = n|\mathcal{C}_n|$, and the related generating function is $C^{\bullet}(z) = zC'(z)$. Hence, if $\mathcal{C}$ has a singular exponent of $-\alpha$, then $\mathcal{C}^{\bullet}$ has a singular exponent of $-\alpha - 1$.

Let now suppose $\mathcal{C}$ to be a combinatorial class with singular exponent $-\alpha > 0$. By marking $\lceil \alpha \rceil$ atoms in the structures into consideration we obtain a flat size distribution over the marked structures. Thus both Theorems 3.1 and 3.2 can be applied on those marked structures. Once an object is generated, we erase its marks, and then obtain an object of $\mathcal{C}$ uniformly at random among objects of the same size.

Note that differentiating a specification is costly in the size of the initial specification but this size growth does not interfere our complexity measure.

4.2 Singular Boltzmann sampler

Another method to deal with the specific case when $0 < -\alpha < 1$ consists in avoiding the computation of the appropriate parameter by choosing the singularity as parameter, since the size distribution has a heavy tail. Note, in this context of singular sampling, the average size of a generated object is infinite. Thus we rely on anticipated rejection to sample an object in a reasonable time.

Theorem 4.1. *Let $\mathcal{C}$ be a combinatorial class whose generating function is Δ-singular with an exponent $0 < -\alpha < 1$.Then the cumulated size T_n of the objects generated and rejected by an approximate Boltzmann sampler with anticipated rejection, $\nu C(\rho, n, \varepsilon)$, satisfies, when n tends to infinity:*

$$\mathbb{E}(T_n) \sim n\kappa_s(\varepsilon, \alpha), \qquad \text{where } \kappa_s(\varepsilon, \alpha) = \frac{-\alpha \cdot \left(\frac{(1-\varepsilon)^{\alpha+1}}{\alpha+1} + \frac{(1+\varepsilon)^{\alpha+1})}{-\alpha}\right)}{(1-\varepsilon)^{\alpha} - (1+\varepsilon)^{\alpha}}.$$

Proof. In order to use the Lemma 2.1, we need the three following estimations obtained by Euler-MacLaurin summations on the coefficients of C.

$$\rho C'^{<n(1-\varepsilon)}(\rho) \sim \frac{c_0 n^{\alpha+1}(1-\varepsilon)^{\alpha+1}}{(\alpha+1)\Gamma(\alpha)}, \qquad C^{>n(1+\varepsilon)}(\rho) \sim -\frac{c_0 n^{\alpha}(1+\varepsilon)^{\alpha}}{\alpha\Gamma(\alpha)},$$

and

$$C^{[n(1-\varepsilon),n(1+\varepsilon)]}(\rho) \sim \frac{c_0 n^{\alpha}}{\Gamma(\alpha)} \frac{(1+\varepsilon)^{\alpha} - (1-\varepsilon)^{\alpha}}{\alpha}.$$

□

4.3 Basic Boltzmann sampler

In the case when $-\alpha > 0$, we can still compute an approximation of the parameter x_n solution of $\mathbb{E}_{x_n}(N) = x_n$. However, to obtain this approximation, the computations are much more technical and the time complexities of the basic Boltzmann samplings (with and without anticipated rejection) are worse than the previous cases. Thus we will not give more details on these cases.

5 Comparison of the methods for peaked distributions

The goal of this section is to be able to compare the methods of sampling. An easy approach (that has been already used in [DFLS04]) is to let ε tend to 0 (with some attention because of the two consecutive limits), and then to observe the behaviour of the complexity.

5.1 A general class of objects with $0 < -\alpha < 1$

We are interested in objects of a combinatorial class, whose singular exponent $-\alpha$ belongs to $]0, 1[$. By pointing one atom, we can apply the basic Boltzmann samplers (Theorems 3.1 and 3.2) with the coefficient $-\alpha - 1$ or we can directly apply the singular Boltzmann sampler. We synthesize the results in the following Figure 1

	Exact complexity	Approx. complexity ($\varepsilon \to 0$ and $\frac{1}{\varepsilon} = o(n)$)
Pointed B. s. without anticipated rej.	$n \cdot \frac{\int_{w=0}^{1-\varepsilon} w^{\alpha+1} e^{(-\alpha-1)w} dw + \int_{w=1+\varepsilon}^{\infty} w^{\alpha+1} e^{(-\alpha-1)w} dw}{\int_{w=1-\varepsilon}^{1+\varepsilon} w^{\alpha} e^{(-\alpha-1)w} dw}$	$\frac{n}{2\varepsilon} \left(\left(\frac{e}{\alpha+1}\right)^{\alpha+1} \cdot \Gamma(\alpha+1) \right) + o\left(\frac{n}{\varepsilon}\right)$
Pointed B. s. with anticipated rej.	$n \cdot \frac{\int_{w=0}^{1-\varepsilon} w^{\alpha+1} e^{(-\alpha-1)w} dw + \int_{w=1+\varepsilon}^{\infty} (1+\varepsilon)\ w^{\alpha} e^{(-\alpha-1)w} dw}{\int_{w=1-\varepsilon}^{1+\varepsilon} w^{\alpha} e^{(-\alpha-1)w} dw}$	$\frac{n}{2\varepsilon} \left(\left(\frac{e}{\alpha+1}\right)^{\alpha+1} \cdot \Gamma(\alpha+1) - \frac{1}{\alpha+1} \right) + o\left(\frac{n}{\varepsilon}\right)$
Singular B. s.	$-\alpha n \left(\frac{(1-\varepsilon)^{\alpha+1}}{\alpha+1} + \frac{(1+\varepsilon)^{\alpha+1}}{-\alpha} \right) \left((1-\varepsilon)^{\alpha} - (1+\varepsilon)^{\alpha} \right)^{-1}$	$\frac{n}{2\varepsilon} \cdot \frac{1}{-\alpha(\alpha+1)} + o\left(\frac{n}{\varepsilon}\right)$

Figure 1: *Asymptotics of the average cumulated size of rejected objects, when* $0 < -\alpha < 1$.

In the light of this study, we can see that the different optimization for Boltzmann sampling proposed, in the original paper [DFLS04], are efficient, even if the gained factor is only linear. If we are interested in exact Boltzmann sampling, i.e. generating objects of an exact given size, then the previous results are valid for $\varepsilon = \frac{1}{2n}$. In this context, we obtain quadratic complexity samplers.

5.2 Aperiodic and strongly connected grammars, $-\alpha = 1/2$

Let us first recall that in this case, the approximate complexity of the Boltzmann singular methods was already presented in [DFLS04].

By taking $-\alpha = 1/2$ in the previous formulas (when ε tends to 0 but with $\frac{1}{\varepsilon} = o(n)$), we obtain the complexity $(\sqrt{\frac{\pi e}{2}} - 1)\frac{n}{\varepsilon} + O(n)$ (with $\sqrt{\frac{\pi e}{2}} - 1 \approx 1.066$) in the case of the pointing sampling with anticipated rejection. In the case of singular sampling, the complexity equals to $2 \cdot \frac{n}{\varepsilon} + O(n)$. Note that the pointing sampling without anticipated rejection is roughly equivalent to the one of the singular sampling. But, pointing sampling with anticipated rejection is almost twice as good!

References

[BBJ14] A. Bacher, O. Bodini, and A. Jacquot. Efficient random sampling of binary and unary-binary trees via holonomic equations. *ArXiv e-prints*, 2014.

[BFP10] O. Bodini, E. Fusy, and C. Pivoteau. Random sampling of plane partitions. *Combinatorics, Probability & Computing*, 19(2):201–226, 2010.

[BLR13] O. Bodini, J. Lumbroso, and N. Rolin. Analytic samplers and the combinatorial rejection method. *CoRR*, abs/1304.1881, 2013.

[BN07] F. Bassino and C. Nicaud. Enumeration and random generation of accessible automata. *Theor. Comput. Sci.*, 381(1-3):86–104, 2007.

[BNW08] F. Bassino, C. Nicaud, and P. Weil. Random generation of finitely generated subgroups of a free group. *IJAC*, 18(2):375–405, 2008.

[BP10] O. Bodini and Y. Ponty. Multi-dimensional boltzmann sampling of context-free languages. *Analco'10*, 2010.

[BRS12] O. Bodini, O. Roussel, and M. Soria. Boltzmann samplers for first-order differential specifications. *Discrete Applied Mathematics*, 160(18):2563–2572, 2012.

[CD09] B. Canou and A. Darrasse. Fast and sound random generation for automated testing and benchmarking in objective caml. *Workshop on ML*, 2009.

[DFLS04] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability & Computing*, 13(4-5):577–625, 2004.

[Duc11] P. Duchon. Random generation of combinatorial structures: Boltzmann samplers and beyond. In *Winter Simulation Conference*, pages 120–132, 2011.

[FFP07] P. Flajolet, E. Fusy, and C. Pivoteau. Boltzmann sampling of unlabeled structures. In *ANALCO*, pages 201–211, 2007.

[FS09] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge Univ. Press, 2009.

[FZC94] P. Flajolet, P. Zimmermann, and B. Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theor. Comput. Sci.*, 132(2):1–35, 1994.

[NW78] A. Nijenhuis and H. S. Wilf. *Combinatorial algorithms for computers and calculators.* Computer science and applied mathematics. Academic Press, New York, 1978.

[PSS12] C. Pivoteau, B. Salvy, and M. Soria. Algorithms for combinatorial structures: Well-founded systems and Newton iterations. *J. of Combinatorial Theory, Series A*, 119:1711–1773, 2012.

[Rém85] J. L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application a leur génération aléatoire. *ITA*, 19(2):179–195, 1985.