# Chaos and Monte-Carlo approximations of the Flip-Annihilation process

**A. D. Ramos**[1] **A. Toom**[2]


## Abstract

The flip-annihilation process is a random particle process with one-dimensional local interaction in discrete time, initially presented by one of us, namely Toom in 2004. Its components are enumerated by integer numbers and every component has two states, "minus" and "plus". At every time step two transformations occur. The first one, called "flip", independently turns every minus into plus with probability $\beta$. The second one, called "annihilation", acts thus: whenever a plus is a left neighbor of a minus, both disappear with probability $\alpha$ independently from other components. What is interesting about this process is that it is ergodic for $\beta > \alpha/2$ and non-ergodic for $\beta < \alpha^2/250$. It is natural to conjecture that there is some transition curve, which we call the *true curve* and denote by $\beta = \mathsf{true}(\alpha)$, which separates the areas of ergodicity and non-ergodicity of this process from each other. The estimates, mentioned above, albeit rigorous, leave a large gap between them and the present article's purpose is to obtain some closer, albeit non-rigorous, approximations of the true curve. We do it in two ways, one of which is a chaos approximation and the other is a Monte Carlo simulation. Thus we obtain two curves, which are much closer to each other than the rigorous estimations. Also we fill in, albeit only numerically, another shortcoming of the rigorous estimation $\beta < \alpha^2/250$, namely that it leaves us uncertain whether the true curve has a zero or positive slope at the point $\alpha = \beta = 0$. Both approximate curves have a positive slope at $\alpha = 0$, as we hoped.


**Key words:** Cellular automata; particle random process; variable length; Monte Carlo; chaos approximation, mean-field approximation.

[1]University of Pernambuco, Polytechnic school of Pernambuco, Recife, PE, Brazil.
E-mail: alexramos@upe.poli.br

[2]Federal University of Pernambuco, Department of Statistics, Recife, PE, 50740-540, Brazil.
E-mail: toom@de.ufpe.br, toom@member.ams.org

# 1 Introduction

For most interacting particle systems, considered till now, the set of components, also called the space, does not change during the process. We call such operators and processes *constant-length*. We call processes, where the sites themselves may appear or disappear, *varible-length* processes. There are quite a few studies of them now, including Toom's [6, 7, 8]. Maes' recent survey [1] refers to a similar kind of processes as processes with *complex architecture* as part of his discussion of new trends in interacting particle systems. Another study of similar processes is presented in Malyshev's works [2, 3], which are motivated by some connections between computer science and quantum gravity.

Let $\mathbb{Z}$ be the set of integer numbers and $\mathcal{A}$ a finite set called *alphabet*. We call $\mathcal{A}^{\mathbb{Z}}$ the *configuration space*. Any $s \in \mathcal{A}^{\mathbb{Z}}$ is called a *configuration*, which may be denoted by $s = (s_i)$, where $s_i$ denotes its *component* at the position $i \in \mathbb{Z}$. We call elements of $\mathcal{A}$ *letters*. We call any finite sequence of letters a *word*. We denote by $\mathcal{A}^n$ the set of $n$-letter words. We call a *thin cylinder* any set of the form

$$\mathcal{C} = \{s \in \mathcal{A}^{\mathbb{Z}} : s_i = a_i \text{ for all } i \in [m, n]\}, \text{ where } a_i \in \mathcal{A}. \tag{1}$$

We consider only normalized measures on $\mathcal{A}^{\mathbb{Z}}$, that is on the $\sigma$-algebra generated by thin cylinders (1). We call a measure on $\mathcal{A}^{\mathbb{Z}}$ *uniform* if it is invariant under translations along $\mathbb{Z}$. For any uniform measure $\mu$ and any word $W = (a_1, a_2, \ldots, a_n) \in \mathcal{A}^n$ we may write

$$\mu(W) = \mu(a_1, a_2, \ldots, a_n) = \mu(s_{i+1} = a_1, s_{i+2} = a_2, \ldots, s_{i+n} = a_n). \tag{2}$$

Since the right side of (2) does not depend of $i$, we may use the left side and call it the *frequency* of the word $W$ in the uniform measure $\mu$. We denote by $\mathcal{M}$ the set of normalized measures on $\mathcal{A}^{\mathbb{Z}}$ and by $\mathcal{M}_u$ the set of normalizad uniform

measures on $\mathcal{A}^{\mathbb{Z}}$.

Random processes with discrete time are usually defined by some operator $P$, which transforms any normalized measure into another normalized measure. By a process we mean a sequence of measures $\mu,\ \mu P,\ \mu P^2, \ldots,$ where $\mu P^n$ is the result of $n$ applications of an operator $P$ to the initial measure $\mu$. (We write operators on the right side of measures on which they act.) A measure $\mu$ is called *invariant* for an operator $P$ if $\mu = \mu P$. Since we deal mostly with uniform measures, we use the following definition of ergodicity. An operator $P : \mathcal{M}_u \to \mathcal{M}_u$ is called *ergodic* if the limit of $\mu P^t$ exists (on the algebra generated by thin cylinders) and is one and the same for all initial measures $\mu \in \mathcal{M}_u$.

Now let us concentrate on the process introduced in [7], which we call *flip-annihilation.* Let us describe it briefly. The components are enumerated by integer numbers. Each component can assume only two states $\oplus$ and $\ominus$, called *plus* and *minus* respectively. Our operator is a superposition of the following two operators. The operator *flip* denoted by $\mathsf{Flip}_\beta$ is very simple. Under its action every minus turns into plus with a probability $\beta$ independently from other components. The *annihilation* operator is denoted by $\mathsf{Ann}_\alpha$. Under its action, whenever a pair $(\oplus,\ \ominus)$ occurs in a configuration, it disappears with probability $\alpha$ independently of what happens at other places. In other words, when a plus and a minus ocuppy the $i$-th and $(i+1)$-th sites respectively, both are eliminated with probability $\alpha$ and the components ocuppying the $(i-1)$-th and $(i+2)$-th sites become neighbors. With the complementary probability $1-\alpha$ these components remain unchanged.

By the flip-annihilation process we mean the superposition $\mathsf{Flip}_\beta \mathsf{Ann}_\alpha$ applied iteratively to an initial measure. (At every step first acts the operator $\mathsf{Flip}_\beta$, then $\mathsf{Ann}_\alpha$.) This article is a numerical study of this process.

We denote by $\delta_\ominus$ and $\delta_\oplus$ the measures concentrated in the configurations *"all minuses"* and *"all pluses"* respectively. Of course, $\delta_\oplus\left(\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha\right)=\delta_\oplus$; so, if we start our process with all pluses, we remain there forever. Therefore our operator $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ is ergodic if and only if

$$\forall\ \mu\in\mathcal{M}_u:\lim_{t\to\infty}\mu\left(\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha\right)^t=\delta_\oplus. \tag{3}$$

The following theorem sums up the most relevant here of what we know about ergodicity of the operator $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$.

**Theorem 1.**

a) if $\beta>\alpha/2$, then $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ is ergodic.

b) if $\beta<\alpha^2/250$, then $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ is not ergodic and has at least two different invariant measures.

For the case $\alpha<1$ this theorem was proved in [5, 7, 8]. For the case $\alpha=1$ it is explained in [5] and a complete proof is available in [4]. Theorem 1 is illustrated by two curves $\beta=\alpha/2$ and $\beta=\alpha^2/250$ on figure 2. We call these curves the *rigorous estimations*. We use a logarithmic scale on the vertical axis in this figure because in the usual scale the curve $\beta=\alpha^2/250$ would be too close to the horizontal axis.

In addition to the facts, which are already proved, we conjecture that whenever $\beta$ increases, the operator $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ cannot pass from ergodicity to non-ergodicity. Under this assumption, for every $\alpha\in[0,1]$ there is a value of $\beta$ in $[0,\ 1]$, which we denote by $\mathsf{true}(\alpha)$, such that the operator $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ is ergodic for $\beta>\mathsf{true}(\alpha)$ and non-ergodic for $\beta<\mathsf{true}(\alpha)$. We assume also that the function $\mathsf{true}(\alpha)$ is continuous, which allows us to speak about the curve $\beta=\mathsf{true}(\alpha)$, which serves as the boundary between the regions of ergodicity $\beta>\mathsf{true}(\alpha)$ and non-ergodicity $\beta<\mathsf{true}(\alpha)$. We call the set $\{(\alpha,\beta):\beta=\mathsf{true}(\alpha)\}$ the *true curve*. Our main goal is to approximate it. Of course, the true curve (if it exists) is sandwiched between

the rigorous estimations, but they are pretty far from each other; we want better numerical estimations.

## 2 Chaos approximation

We denote by $\mathcal{C} : \mathcal{M}_u \to \mathcal{M}_u$, the well-known *chaos operator*. Its action amounts to mixing randomly all the components. In other words, for each $\mu \in \mathcal{M}_u$ the measure $\mu\,\mathcal{C}$ is a product-measure with the same frequencies of all the letters as $\mu$ has. (This method is also known as mean-field approximation.) The chaos operator allows us to approximate a given process $\mu\,P^t$ on the configuration space $\mathcal{A}^{\mathbb{Z}}$ by another process $\mu\,(\mathcal{C}\,P)^t$ on the same space. (Every time we apply first $\mathcal{C}$, then $P$.) Thus, instead of the original process, whose set of parameters is infinite or very large, we deal with the evolution of densities of letters, that is a finite and limited set of parameters. Since densities of the letters sum up to one, the number of independent parameters in the chaos approximation equals the number of letters in the alphabet minus one. In our case, with only two letters, we deal with only one parameter: as such we choose the density of pluses.

Due to the properties of the chaos operator, the density of pluses in the measure $\mu\,\mathcal{C}\,\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ depends only on the density of pluses in the measure $\mu$ and this dependence may be expressed by the formula

$$f(x) = \frac{b - \alpha \cdot b(1-b)}{1 - 2\alpha \cdot b(1-b)}, \tag{4}$$

where $x$ denotes the fequency of pluses in the measure $\mu$, $f(x)$ denotes the density of pluses in the measure $\mu\,\mathcal{C}\,\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ and $b = x + (1-x)\beta$. ($b$ is the density of pluses in the measure $\mu\,\mathcal{C}\,\mathsf{Flip}_\beta$.) Since the maximal value of $b(1-b)$ is $1/4$, the denominator of (4) is not less than $1/2$. So $f(t)$ is defined and continuous for $x,\ \alpha,\ \beta \in [0,\ 1]$. Thus the study of the operator $\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ is substituted by

a study of the operator $\mathcal{C}\,\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$, which boils down to the study of the one-dimensional dynamical system $f : [0,1] \to [0,1]$ with parameters $\alpha,\ \beta \in [0,\ 1]$. As usual, we call a *fixed point* of this system a value of $x \in [0,\ 1]$ such that $f(x) = x$. We call our dynamical system *ergodic* if it has a unique fixed point $x_{\text{fixed}}$ and

$$\forall\ x \in [0,\ 1] : \lim_{t\to\infty} f^t(x) = x_{\text{fixed}},$$

where $f^t$ means the $t$-th iteration of $f$.

**Theorem 2.** The chaos approximation $\mathcal{C}\,\mathsf{Flip}_\beta\,\mathsf{Ann}_\alpha$ is ergodic if $\beta > \beta^*(\alpha)$ and is not ergodic if $\beta \le \beta^*(\alpha)$, where

$$\beta^*(\alpha) = \begin{cases} \dfrac{4-\alpha-2\sqrt{4-2\alpha}}{\alpha} & \text{if } \alpha > 0, \\ 0 & \text{if } \alpha = 0. \end{cases} \tag{5}$$

Thus for the chaos approximation we know exactly the curve dividing ergodicity and non-ergodicity: it is the continuous curve $\beta = \beta^*(\alpha)$: it starts at the origin with the slope $1/8$, grows smoothly and reaches $3 - 2\sqrt{2} \approx 0.17$ at $\alpha = 1$. The graph of this curve is labeled "Chaos" in the figure 2.

**Proof of theorem 2.**

First let $\alpha = 0$. It is easy to observe that in this case our dynamical system is ergodic for all $\beta > 0$ and non-ergodic for $\beta = 0$, so theorem 2 is true. Now let $\alpha > 0$. If $\beta$ equals zero or one, theorem 2 is evident. So now we additionally assume that $0 < \beta < 1$. Solving the equation $f(x) = x$ explicitly, we find out that all the fixed points of $f$ are those of the values

$$p_1 = \frac{\alpha - 3\alpha\beta - \sqrt{\Delta}}{4\alpha(1-\beta)}, \qquad p_2 = \frac{\alpha - 3\alpha\beta + \sqrt{\Delta}}{4\alpha(1-\beta)}, \qquad p_3 = 1, \tag{6}$$

which are real and belong to $[0,\ 1]$, where

$$\Delta = \alpha^2\beta^2 + 2\alpha^2\beta + \alpha^2 - 8\alpha\beta. \tag{7}$$

Since $\alpha > 0$ and $\beta < 1$, the denominator of $p_1$ and $p_2$ in (6) is not zero. Since $\beta > 0$,

$$0 < p_1 \leq p_2 < p_3 = 1 \tag{8}$$

whenever $p_1$ and $p_2$ are real, that is $\Delta \geq 0$. $\Delta$ equals zero at $\beta = \beta_1(\alpha)$ and $\beta = \beta_2(\alpha)$, where

$$\beta_1(\alpha) = \frac{4 - \alpha - 2\sqrt{4 - 2\alpha}}{\alpha} \text{ and } \beta_2(\alpha) = \frac{4 - \alpha + 2\sqrt{4 - 2\alpha}}{\alpha}. \tag{9}$$

Here $\beta_1$ is what we called $\beta^*$. According to (7), $\Delta$ is a quadratic function of $\beta$ with a positive second-degree coefficient, so $\Delta$ is negative when $\beta$ is between the roots (9) and positive when $\beta$ is less that $\beta_1(\alpha)$ or greater than $\beta_2(\alpha)$. It is easy to observe that

$$\forall\ \alpha > 0\ :\ 0 < \beta_1(\alpha) < 1 < \beta_2(\alpha). \tag{10}$$

If $\beta < \beta_1(\alpha)$, then $\Delta > 0$ and our dynamical system has three different fixed points $p_1$, $p_2$ and $p_3$, so it is non-ergodic. If $\beta = \beta_1(\alpha)$, then $\Delta = 0$ and our dynamical system has two different fixed points $p_1 = p_2$ and $p_3$, so it is non-ergodic too.

Now let $\beta > \beta_1(\alpha)$. In this case $\Delta < 0$, so $f$ has only one fixed point $p_3 = 1$. Let us prove that for all $x_0 \in [0, 1]$, $f^t(x_0)$ tend to 1 when $t$ tends to infinity. As we know, $f$ is continuous in $[0, 1]$. So, $g(x) = f(x) - x$ is also continuous. It is easy to calculate that $g(0) > 0$. Since $g$ is continuous, equals zero only at the point $x = 1$ and $g(0) > 0$, we conclude that $g(x) > 0$ for all $x < 1$, whence $f(x) > x$ for all $x < 1$. Then for all $x_0 < 1$ the sequence $f^t(x_0)$ is growing and limited by 1 and therefore has a limit, which is a fixed point of $f$. But 1 is the only fixed point of $f$, so $f^t(x_0) \to 1$ when $t \to \infty$. *Theorem 2 is proved completely.*

In fact, we can describe completely the limit behavior of this dynamical system:

$$\text{If } \Delta < 0, \quad \text{then } \lim_{t\to\infty} f^t(x_0) = p_3 = 1 \text{ for all } x_0.$$

$$\text{If } \Delta = 0, \quad \text{then } \lim_{t\to\infty} f^t(x_0) = \begin{cases} p_1 = p_2 & \text{if } x_0 \le p_1 = p_2, \\ p_3 = 1 & \text{if } x_0 > p_1 = p_2. \end{cases}$$

$$\text{If } \Delta > 0, \quad \text{then } \lim_{t\to\infty} f^t(x_0) = \begin{cases} p_1 & \text{if } x_0 < p_2, \\ p_2 & \text{if } x_0 = p_2, \\ p_3 = 1 & \text{if } x_0 > p_2. \end{cases}$$

## 3 Monte Carlo simulation

Along with the process on $\mathcal{A}^{\mathbb{Z}}$ we shall consider its analogs on finite spaces. When our operators act on finite configurations, every act of elimination of $(\oplus, \ominus)$ decreases the length of the configuration by two, whence in the average the number of components decreases and the process degenerates into a finite sequence of pluses (provided $\beta > 0$), which remains one and the same forever. However, the time needed for this may depend drastically on the values of our parameters $\alpha$ and $\beta$.

Let us consider the following finite analog of the flip-annihilation process. It is a Markov chain with a countable set $\Omega$ of states called *circulars*. The circulars are similar to words as they also are finite sequences of pluses $\oplus$ and minuses $\ominus$, but now we imagine these sequences to have circular form. We denote by $|C|$ the number of components in a circular $C$. The indices of these components are remainders modulo $|C|$ (see figure 1 where $|C| = n$.)

(We could use words instead of circulars, but this would necessitate special definitions at the ends when we transform them.) In most of our Monte Carlo

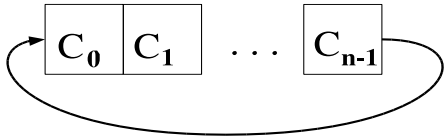


Figure 1: *A circular* $C$ *with* $|C| = n$ .

simulations, the initial circular $C$ consisted of 1000 minuses. In every single experiment, the integer time $t$ grew from zero to at most 100 000. The circular obtained at time $t$ was denoted by $C^t$ and its $i$-th components were denoted by $C_i^t$, where $i = 0, \ldots, |C^t| - 1$.

We say that a word $W = (a_1, a_2, \ldots, a_n)$ *appears* at a place $i$ in a circular $C = (c_1, \ldots, c_n)$ if

$$c_{i+1} = a_1,\ c_{i+2} = a_2, \ldots, c_{i+n} = a_n.$$

We denote by $\mathsf{quant}(W|C)$ the quantity of different places where the word $W$ appears in a circular $C$. After that, we define the *frequency* of $W$ in $C$ as follows:

$$\mathsf{freq}(W|C) = \frac{\mathsf{quant}(W|C)}{|C|}. \tag{11}$$

Let us describe a procedure, which we call *Imitation* and which is a Monte Carlo imitation of our process. This procedure generates a sequence of circulars in the following inductive way.

**Base of induction.** The initial circular $C^0$ consists of 1000 minuses.

**$t$-th induction step.** Given a circilar $C^t$, where $t = 0,\ 1,\ 2, \ldots$ we perform three procedures:

*First procedure imitating the action of flip:* every component of $C^t$, which is a minus, becomes a plus with a probability $\beta$ independently from other components. (In more technical detail, for every minus in $C^t$ we generate a new random variable distributed uniformly in $(0,\ 1)$ and change this minus into plus if this variable is

less than $\beta$.) We denote the resulting circular by $(C')^t$.

*Second procedure imitating the action of annihilation:* whenever a component of $(C')^t$, which is a plus, is a left neighbor of a component, which is a minus, both are eliminated from the circular with a probability $\alpha$ independently from other components. (In technical details, for every such pair we generate a new random variable distributed uniformly in $(0,\ 1)$ and perform this elimination if this variable is less that $\alpha$.) We denote the resulting circular by $(C'')^t$.

*Third procedure which helps to imitate the infinite process:* given $(C'')^t$, we generate a new circular, namely $C^{t+1}$, in the following way: if $|(C'')^t| < N_{min}$, where $N_{min} = 500$, then $C^{t+1}$ is obtained from $(C'')^t$ by concatenating it with its copy and thereby duplicating its length; otherwise $C^{t+1} = (C'')^t$.

**When we stop:** given a constant $T = 100000$, we stop when $t = T$ or there is none minus in the circular $C^t$.

Let us explain why we need the third procedure. Remember that under the action of our operator, components can disappear, but not appear; so for any $\beta > 0$ the length of any finite circular decreases in the average and finally the process degenerates into a circular consisting only of pluses. The third procedure allowed us to postpone this and thereby helped us to make our simulation more similar to the infinity process.

Thus the procedure Imitation is described. We used it for various purposes in our study, but right now we use it for only one purpose: to attribute the appropriate value to a Boolean variable denoted by $E$ (which means ergodicity), namely $E$ is given the value $yes$ if the last circular $C^t$ contains none minus; otherwise $E$ is given the value $no$. If $E = yes$, we interpret this as a suggestion that the process with the given values of $\alpha$ and $\beta$ is ergodic; the result $E = no$ is taken as a suggestion that our process is non-ergodic.

In fact we used Imitation within a cycle with growing $\beta$: we started with $\beta = 0$ and then iteratively performed Imitation and increased $\beta$ by 0.001 and repeated this until $\beta$ reached the value 1 or $E$ got the value $yes$, that is ergodicity was suggested. Thus we obtained a certain value of $\beta$. In fact, we performed this cycle 5 times and recorded the arithmetical average of the 5 values of $\beta$ thus obtained.

Remember that all this was done with a certain value of $\alpha$. In fact we considered 1000 values of $\alpha$, namely the values $\alpha_i = 0,001 \cdot i$ for $i = 1, \ldots, 1000$. The corresponding recorded value of $\beta$ was denoted by $\beta_i$. Thus we obtained 1000 pairs $(\alpha_i,\ \beta_i)$. The graph called M. C. on figure 2 consists of these pairs plotted.

You can see that the Monte Carlo "curve" is not exactly a curve, it is somewhat fuzzy. If instead of five procedures we had more for every value of $\alpha$, this graph would be thinner. However, even such as it is, it gives some idea of the behavior of our process.

We see that the Chaos and M.C. curves are much closer to each other than the rigorous estimations and we conjecture that they are closer to the true curve also.

### 3.1 Estimation of $s(\alpha,\ \beta)$

In [7] a function $s(\alpha,\ \beta)$ was defined as the supremum of density of pluses in the measure $\mu_t$ over all natural $t$. For every $\alpha$ the function $s(\alpha,\ \beta)$ was proved not to be continuous as a function of $\beta$. We wanted to estimate $s(\alpha,\ \beta)$ numerically, but to estimate it directly was difficult, so, instead of that, we estimated

$$\overline{s(\alpha,\ \beta)} = \max\{\mathsf{freq}(\oplus|C^t) : t = 0, \ldots, 100\,000\}.$$

In the area of ergodicity (white area) $\overline{s(\alpha,\ \beta)} = 1$. Figure 3 shows the values of $\overline{s(\alpha,\ \beta)}$ in the other area, where our finite approximation suggests non-ergodicity.

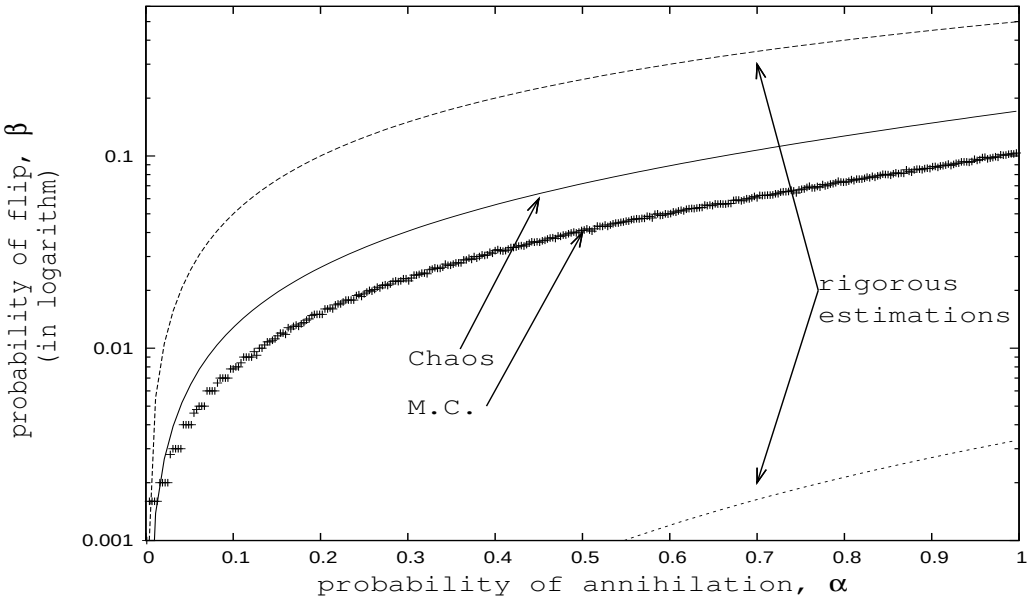


Figure 2: *This graph shows both rigorous estimations and the two approximations of* $\mathsf{true}(\alpha)$ *: the chaos approximation(Chaos) and the Monte Carlo approximation (M. C.). Every point of the latter curve was obtained as an average of* 5 *independent experiments.*

The values of $\overline{s(\alpha,\ \beta)}$, are represented by colors according to the rule shown in the color box on the right side. All the values of $\overline{s(\alpha,\ \beta)}$, which we obtained for all the non-ergodic area, were less or equal to $0.14$, which illustrates non-continuity of $s(\alpha,\ \beta)$ as a function of $\beta$.

We studied in more detail the behavior of $\overline{s(\alpha,\ \beta)}$ near the critical curve: we took a few values of $\alpha$ and for everyone of them made 100 experiments, in each of them making $\beta$ grow from zero to one with an increment $0.001$, all the time calculating the supremum of density of pluses, which we denoted by $\overline{s_i(\alpha,\ \beta)}$. Then we defined

$$E[s(\alpha,\ \beta)] = \frac{1}{100}\sum_{i=1}^{100} \overline{s_i(\alpha,\ \beta)} \tag{12}$$

and used (12) as an approximation for $s(\alpha,\ \beta)$. The behavior of $E[s(\alpha,\beta)]$ for $\alpha = 0.25,\ 0.5$ and $0.75$ is shown in figure 4. We observe that when $\beta$ increases, $E[s(\alpha,\beta)]$ also increases and that near the $\mathsf{true}(\alpha)$ there is an abrupt increase. For the same sample we evaluated the variance. The error bars in figure 4 show

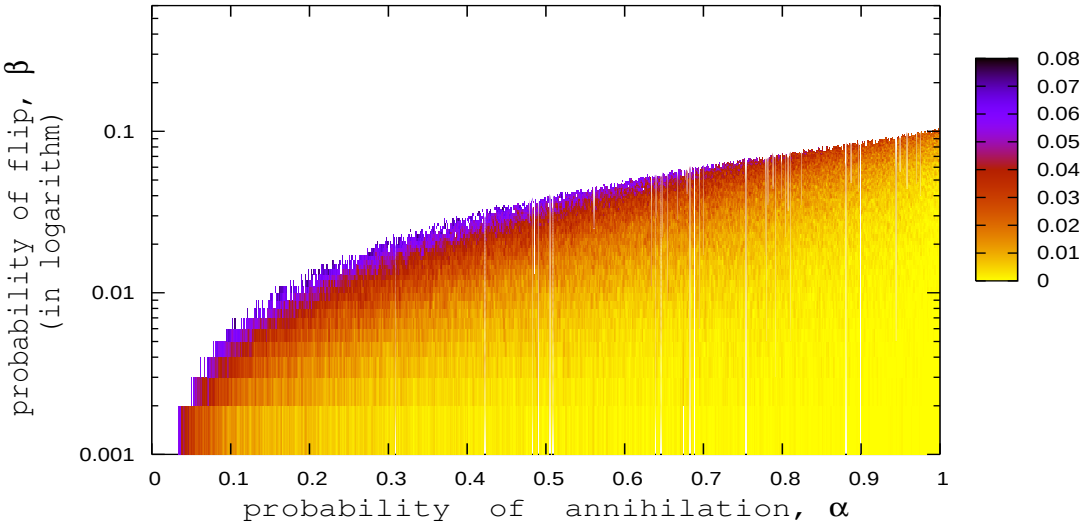


Figure 3: *Here we used colors to represent the values of* $\overline{s(\alpha,\ \beta)}$ *in the area, where the process is suggested to be non-ergodic. The color box on the right side shows how colors from yellow to black represent the values of* $\overline{s(\alpha,\ \beta)}$. *For better visualization, we excluded the values greater than* 0.08, *which constitute less than* 1% *of all data.*

the variance of those $\overline{s_i(\alpha,\ \beta)}$ which we used to obtain $E[s(\alpha,\ \beta)]$ in (12). We see that the variance is the largest near the critical value $\mathsf{true}(\alpha)$. We notice similar qualitative behavior of $E[s(\alpha,\beta)]$ for all the three values of $\alpha$.

To study the first order phase transition of $s(\alpha,\beta)$ in more detail, we studied the behavior of $\overline{s(0.5,\beta)}$, taking the initial circulars $C$ consisting of $|C| = 125$, 250, 500, 1000, 2000 and 4000 minuses. For each one of them, the maximal time was taken $100 \cdot |C|$. Figure 5 shows the results of this experiment. In this case we restricted our attention to a window of observation, where, as we believed, was the critical value $\mathsf{true}(0.5)$, which separates the area, where $\overline{s(0.5,\beta)} = 1$, from the area, where $\overline{s(0.5,\beta)} < 1$.

We denoted by $a_{|C|}$ the minimum of those values of $\beta$, for which the variance of $E[s(0.5,\beta)]$ was greater than 0.01 and by $b_{|C|}$ the maximum of those values $\beta$, for which the variance of $E[s(0.5,\beta)]$ was greater than $0,01$. Using the behaviour

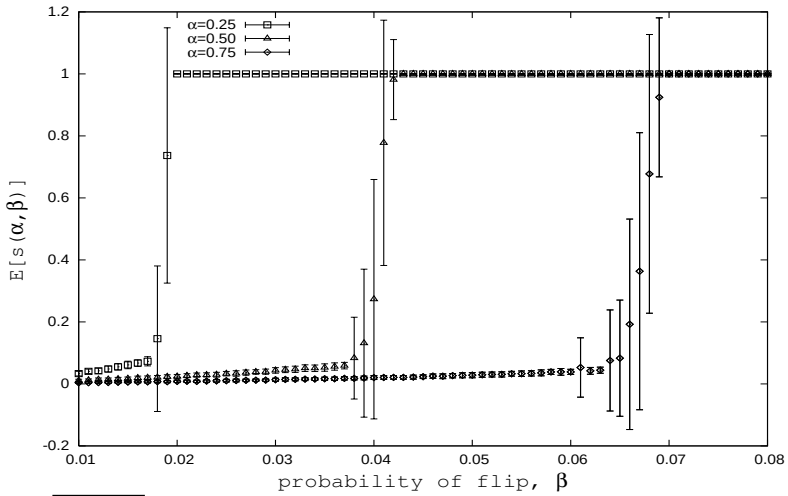


Figure 4: *Behavior of* $\overline{s(\alpha,\beta)}$ *for three values of* $\alpha$, *namely* $\alpha = 0.25,\ 0.5$ *and* $0.75$. *In each case it grows sharply near the critical value. For each value of* $\beta$ *we made* $100$ *independent experiments. Error bars represent the standart deviation.*

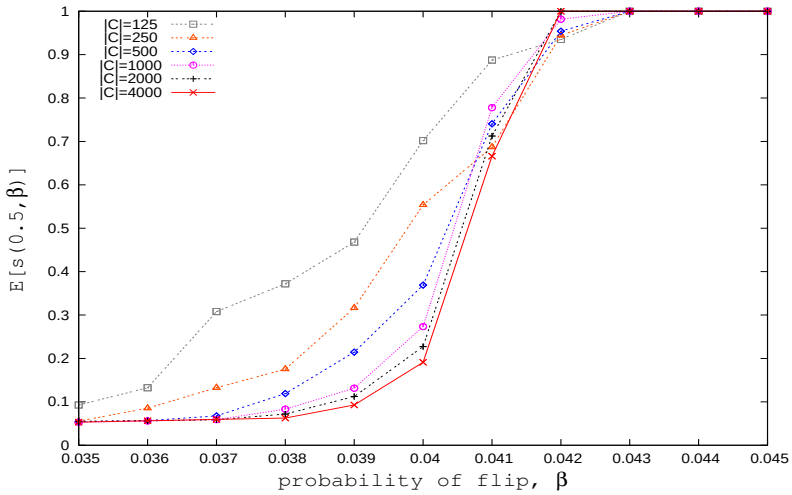


Figure 5: *We recorded values of* $\overline{s(0.5,\beta)}$ *for different lengths of initial circulars. For each* $\beta$ *we conducted 100 experiments.*

of variance near $\mathsf{true}(\alpha)$, we estimated the interval $[a_{|C|},\ b_{|C|}]$, which contains all the values of $\beta$, whose variance was greater than $0.01$. The results are shown in table 1.

| $\lvert C\rvert$ | $[a_{\lvert C\rvert}, b_{\lvert C\rvert}]$ | $E[s(0.5, 0.001)]$ | $Var[E[s(0.5, 0.001)]]$ |
|---|---|---|---|
| 125 | $[0.007, 0.042]$ | 0.0156 | 0.00411 |
| 250 | $[0.014, 0.042]$ | 0.0121 | 0.00252 |
| 500 | $[0.023, 0.042]$ | 0.0106 | 0.00179 |
| 1000 | $[0.030, 0.042]$ | 0.00076 | 0.00106 |
| 2000 | $[0.038, 0.042]$ | 0.00087 | 0.00085 |
| 4000 | $[0.039, 0.041]$ | 0.00095 | 0.00062 |

Table 1: *This table shows* $E[s(0.5, \beta)]$ *and the correspondent* $[a_{|C|}, b_{|C|}]$ *when* $|C|$ *increases.*

Table 1 shows that when $|C|$ increases, the length of the interval $[a_{|C|}, b_{|C|}]$ decreases. Also, we observed that $E[s(0.5, \beta)]$ shows a behaviour more stable when $|C|$ increases since its variance is decreasing. We conjecture that when $|C|$ tends to infinity, the interval $[a_{|C|}, b_{|C|}]$ degenerates into a point and this point is the value of $\mathsf{true}(\alpha)$. Moreover, we conjecture that when $|C|$ tends to infinity, our estimation (12) for all values of $\beta$ converges to the supremum of density of pluses in $\mu_t$, that is to $s(\alpha, \beta)$.

We also studied how the results of our M.C. approximation depend on the parameter $N_{min}$. This suggests that our conclusions about phase transition and slope remain valid for others values of $N_{min}$. This confirms our belief that this computational model approximates well the infinity process.

### 3.2 The slope at zero

It is natural to expect our process to converge to a process with continuous time when $\alpha$ and $\beta$ tend to zero, their proportion remaining constant. If this is true, the true curve should have a positive slope at $\alpha = \beta = 0$, but our rigorous estimations do not imply this. For this reason we are interested in the value of this

slope for our approximations. We have already seen that the chaos approximation has a positive slope at the origin, namely $1/8$. Now let us evaluate this slope for the Monte Carlo simulation. The procedure is similar to that we used to obtain the separating curve shown in figure 2, but with the following modifications: for every $j \in \{0, 1, 2, 3, 4, 5\}$, we take $\alpha$ varying from zero to $1/2^j$ with an increment of $1/(2^j \times 1000)$. So, for every considered value of $j$, we recorded pairs $\{\alpha_i^j,\ \beta_i^j\}$ for $i = 1, \ldots, 1000$. For these pairs we calculated the best fits, linear and quadratic, which we denote by

$$f_L^j(\alpha) = a_j \cdot \alpha \qquad \text{and} \qquad f_Q^j(\alpha) = b_j \cdot \alpha + c_j \cdot \alpha^2.$$

To obtain these fits, we used the least square method. We introduced two functions, $f_L^j$ and $f_Q^j$ which are the "best" fits for the set of datas. These functions are obtained by minimization of

$$D_{f_L}^j = \sum_{i=1}^{1000} (\beta_i - f_L^j(\alpha_i^j))^2 \quad \text{and} \quad D_{f_Q}^j = \sum_{i=1}^{1000} (\beta_i - f_Q^j(\alpha_i^j))^2.$$

| $j$ | $a_j$ | $b_j$ | $c_j$ |
|---|---|---|---|
| 0 | 0.0911 | 0.0632 | 0.0371 |
| 1 | 0.0779 | 0.0718 | 0.0165 |
| 2 | 0.0742 | 0.0737 | 0.0027 |
| 3 | 0.0729 | 0.0738 | −0.0095 |
| 4 | 0.0730 | 0.0745 | −0.0321 |
| 5 | 0.0729 | 0.0712 | 0.0715 |

Table 2: *The first column contains the coefficients of the linear fit, which show convergence to* $\approx 0.073$. *The second and third columns contain the coefficients of the first and second degree terms of the quadratic fit respectively. The second column behaves like the first one, only looser. The third column shows no clear pattern, which suggests that the quadratic term is irrelevant.*

Table 2 shows the coefficients of the linear and quadratic fits. For $\alpha$ near zero the sequence of coefficients $a_j$ of the linear fit, shown in the first column of the table 2, stabilizes around $0.073$ when $j > 2$. The second and third columns show the coefficients of the quadratics fit. The coefficient $b_j$ of the linear term of this fit

fluctuates around the same value as $a_j$, but with less persistence. The coefficient $c_j$ shows no clear pattern of behavior, which suggests that it is irrelevant. We see that the linear fit approximates our curve near the origin $(0, 0)$ at least as well as the quadratic fit. This suggests that the curve is approximatedly linear near the origin with the slope approximated by the values of $a_j$. We conclude that the M.C. approximation has a positive slope at the origin, approximatedly equal $0.073$.

## 4 Conclusions

This work is a numerical study of a variable-length process, which has been proved elsewhere to behave in a non-trivial way. We obtained two approximations of the hypothetical curve $\mathsf{true}(\alpha)$, separating areas of ergodic and non-ergodic behavior of this process, using a chaos approximation and a Monte Carlo simulation. Also, we have presented a computational model to approximate this process, which may be used to approximate other variable-length processes. Finally, we have shown numerically that the slope of the curve $\mathsf{true}(\alpha)$ is positive when the probabilities $\alpha$ and $\beta$ are near zero. Our numerical work may indicate appropriate directions for future research.

## Acknowledgments

We thank the referee for useful correction. We cordially thank D. M. Pianto for a useful discussion of table 1. A. D. Ramos cordially thanks the CNPq/MCT and CT-Info, which supported this work. A. Toom was partially supported by CNPq and PRONEX.

# References

[1] C. Maes. *New Trends in Interacting Particle Systems.* Markov Proc. Rel. Fields. vol. 11 (2005), n. 2, pp. 283-288.

[2] V. Malyshev. *Quantum Evolution of Words.* Theoretical Computer Science, 2002. Avaliable at `http://www-rocq.inria.fr/\%7Emalyshev/Malyshev/papers.htm`.

[3] V. Malyshev. *Quantum Grammars. Part 4.1: KMS states on Quantum Grammars.* Preprint INRIA, 1999, no. 3702, pp. 1-20 and Journal of Mathem. Physics, v. 41 (2000), n. 7, pp. 4508-4520.

[4] A. D. Ramos. *Particle process with variable length.* Ph.D. Thesis. Federal University of Pernambuco, Departament of Statistics, Recife, Pernambuco, Brazil, 2007. (In Portuguese with an abstract in English). Avaliable at `http://de.ufpe.br/~toom/ensino/doutorado/alunos/index.htm`.

[5] A. D. Ramos and A. Toom. *An error correction. Letter to the editor.* Journal of Statistical Physics, volume 131, number 1, april 2008, pp 167-168.

[6] A. Toom. *Particle systems with variable length.* Bulletin of the Brazilian Mathematical Society, v. 33, n. 3, november 2002, pp. 419-425.

[7] A. Toom. *Non-Ergodicity in a 1-D Particle Process with variable Length.* Journal of Statistical Physics, vol. 115, nn. 3/4, 2004, pp. 895-924.

[8] A. Toom. *Every continuous operator has an invariant measure.* Journal of Statistical Physics, vol. 129, n. 2007, pp. 555-566.