

## Short-ranged memory model with preferential growth

Ana L. Schaigorodsky,<sup>\*</sup> Juan I. Perotti,<sup>†</sup> Nahuel Almeida,<sup>‡</sup> and Orlando V. Billoni<sup>§</sup>

*Instituto de Física Enrique Gaviola (IFEG-CONICET), Ciudad Universitaria, 5000 Córdoba, Argentina  
and Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba, 5000 Córdoba, Argentina*



(Received 28 November 2017; published 20 February 2018)

In this work we introduce a variant of the Yule-Simon model for preferential growth by incorporating a finite kernel to model the effects of bounded memory. We characterize the properties of the model combining analytical arguments with extensive numerical simulations. In particular, we analyze the lifetime and popularity distributions by mapping the model dynamics to corresponding Markov chains and branching processes, respectively. These distributions follow power laws with well-defined exponents that are within the range of the empirical data reported in ecologies. Interestingly, by varying the innovation rate, this simple out-of-equilibrium model exhibits many of the characteristics of a continuous phase transition and, around the critical point, it generates time series with power-law popularity, lifetime and interevent time distributions, and nontrivial temporal correlations, such as a bursty dynamics in analogy with the activity of solar flares. Our results suggest that an appropriate balance between innovation and oblivion rates could provide an explanatory framework for many of the properties commonly observed in many complex systems.

DOI: [10.1103/PhysRevE.97.022132](https://doi.org/10.1103/PhysRevE.97.022132)

### I. INTRODUCTION

Usually known as Zipf's law or Pareto distributions, power laws and long-tailed distributions are scattered among many natural systems [1]. For instance, these distributions can be found in physics, biology, earth and planetary sciences, economics and finance, and demography and the social sciences. A short list of the systems where these laws are observed includes the sizes of earthquakes [2] and solar flares [3], the frequency of use of words in human languages [4], the number of species in biological taxa [5], people's annual incomes [6], and the distribution of the popularities of game lines in chess [7].

Several mechanisms have been proposed for generating highly skewed distributions with long tails [1,8,9]. In particular, the underlying probabilistic mechanism of the Yule-Simon process [10,11] is one of the most studied due to its simplicity. The Yule-Simon model, originally inspired by observations of the statistics of biological taxa, reemerged in the literature several times. In fact, its most recent variant, known as preferential attachment, became one of the most important ideas in the early development of the theory of complex networks [12]. Yule's model was introduced to explain the distribution of the number of species in a genus, family, or other taxonomic group. Later, a more particular version of the model was introduced by Simon to explain the frequency of words in literary corpuses, i.e., to explain Zipf's law. Within the context of text generation, the Yule-Simon's model works as follows. At each discrete time step a new word is generated with probability  $p$ , or an already existing word is copied with

probability  $1 - p$ . This process generates a text with few words that are very popular and many others that are barely used. Namely, the popularity distribution of the words generated by this process is a skewed distribution with a power-law tail of exponent  $\alpha = 2 + p/(1 - p)$ , meaning that the number  $n_s$  of words with popularity  $s \gg 1$  satisfy  $n_s \sim s^{-\alpha}$ .

Another interesting effect that is frequently observed in empirical systems that display a long-tailed frequency distribution is the presence of memory in the form of long-range correlations [13–16] or a bursty dynamics [17,18]. These memory effects are not accounted by the standard preferential growth models [7,10,11,19]. To address the observed memory effects, Cattuto, Loreto, and Severdio [20] introduced a variant to the Yule-Simon model incorporating a fat-tailed memory kernel, from now on we call this variant CLS model. The series generated by CLS model exhibit temporal long-range correlations, which are missing in the Yule-Simon model [21,22]. Moreover, it preserves the long-tailed popularity distribution and is able to reproduce quite well the statistics of tag occurrences in the collaborative activity of web users [21]. In a recent work [22], we have shown that the long-range correlations introduced by CLS model allow to describe the memory effects observed in the growing process of an extensive chess database. However, this model fails to describe the interevent time distribution in the use of popular game lines by individual players. In other words, CLS model introduces long-range correlations but fails to produce a *bursty* interevent time dynamics.

Aiming to generate a bursty dynamics, in this work we replace the fat-tailed kernel of CLS model by a finite-sized memory kernel. This modification introduces bursts in the occurrence of sets of words but eliminates the long-range correlations, although its characteristic length exhibits a nontrivial finite-size scaling. It also preserves the power-law distribution expected in Yule-Simon model. However, the distribution exponent is neither the corresponding for the Yule-Simon

<sup>\*</sup>schaigorodsky@famaf.unc.edu.ar

<sup>†</sup>juanpool@gmail.com

<sup>‡</sup>nalmeira@famaf.unc.edu.ar

<sup>§</sup>billoni@famaf.unc.edu.ar

model nor CLS model. It is worth to remark that, in spite of the large amount of works dealing with the mechanism of preferential attachment and its main variations [23–26], which include nonlinear dependence of the kernel [27], fitness [28], and aging effects [20,29–31], to the best of our knowledge, this is the first time that the effect of a finite-sized kernel in the Yule-Simon model is studied.

The article is organized in the following manner. In Sec. I, we introduce the model with the finite-size kernel and analyze the emerging lifetime and popularity distributions. In Sec. II, we characterize the memory effects, such as correlations and burstiness. In Sec. III, we perform an analysis of the kernel state as a function of the model parameters to understand the phenomenology of the model, where we establish an analogy with the statistical mechanics theory of equilibrium system. Finally, in Sec. IV, we provide a discussion of the main results of this work.

## II. MODELS AND POPULARITY DISTRIBUTION ANALYSIS

### A. Models

The mechanism to generate artificial sequences in preferential growth models is as follows. The process begins with an initial state of  $n_0$  randomly generated words, or more generally, elements of different classes. At each discrete temporal step  $t$  there are two options: (i) to introduce an element of a new class with probability  $p$  or (ii) to copy an already existing element with probability  $\bar{p} = 1 - p$ . The way to select the element to be copied in option (ii) establishes the difference between Yule-Simon and CLS model. In the Yule-Simon model every existing element has the same probability to be copied, while in CLS model the probability of copying an element that occurred at time  $t - \Delta t$  is given by the functional form of the memory kernel  $Q(\Delta t)$ ,

$$Q(\Delta t) = \frac{C(t)}{\Delta t + \kappa_C}, \quad (1)$$

where  $C(t)$  is a logarithmic normalization factor and  $\kappa_C$  is the characteristic length or memory kernel extension.

The model we introduce in this work is a simple variant of CLS model. The difference resides in the memory kernel, which in our model is defined by a step function:

$$Q(\Delta t) = \begin{cases} \frac{1}{\kappa} & \Delta t \leq \kappa \\ 0 & \Delta t > \kappa \end{cases}, \quad (2)$$

where  $\kappa$  is the memory kernel extension. We call this variant of the process bounded memory preferential growth model (BMPG model). Figure 1 outlines the process for the BMPG model.

### B. Master equation

As a first analysis of our model we describe the dynamics of the elements inside the memory kernel. Notice that, at each time step  $t$ , the last element in the kernel leaves the kernel in the next time step. As a consequence, a class of elements may go extinct if the leaving element is the last of its class within the kernel—e.g., the pink gridded balls in Fig. 1. In this context, it is interesting to study the extinction dynamics

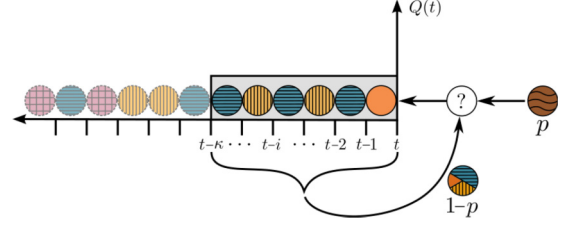


FIG. 1. Outline of BMPG model process depicting how an element is incorporated to the time series at time  $t$ . With probability  $p$  the added element belongs to a new class, i.e., a new color ball is introduced. With complementary probability  $1 - p$  an existing element is copied, chosen uniformly among the last elements of the series. Elements that are more than  $\kappa$  time steps away—the shaded balls in the figure—cannot be copied.

of the elements in the kernel through the analysis of lifetime distributions. Let  $P_n(t)$ , ( $n = 0, 1, \dots, \kappa$ ) be the probability at time  $t$  that a given class of element inside the kernel has popularity  $n$ , and  $\mathbf{P}(t)$  the vector with components  $P_n(t)$  ( $n = 0, 1, 2, \dots, \kappa$ ). The probability  $P_n(t)$  can be approached by a discrete Markov chain process, considering that an element of a class can be copied increasing its popularity  $n \rightarrow n + 1$ , can be removed decreasing its popularity,  $n \rightarrow n - 1$ , or can be copied and removed in the same time step keeping its popularity. In the BMPG process the older element in the kernel is removed. However, to derive the Markov chain we make an approximation, where at each time step we remove instead a randomly chosen element of the kernel. In this approximation we derive the following transition probabilities for a kernel with  $n$  elements of a particular class:

$$\begin{aligned} \Pi_{n,n} &= (1 - p) \left[ \left( \frac{n}{\kappa} \right)^2 + \left( 1 - \frac{n}{\kappa} \right)^2 \right] + p \left( 1 - \frac{n}{\kappa} \right) (n \neq 0), \\ \Pi_{n,n+1} &= (1 - p) \frac{n}{\kappa} \left( 1 - \frac{n}{\kappa} \right), \\ \Pi_{n,n-1} &= (1 - p) \frac{n}{\kappa} \left( 1 - \frac{n}{\kappa} \right) + p \frac{n}{\kappa}, \end{aligned} \quad (3)$$

where  $n = 0, 1, 2, \dots, \kappa$ . Then, if the probability distribution is known at time  $t$ , the probability distribution at time  $t + 1$  can be computed as  $\mathbf{P}(t + 1) = \mathbf{P}(t)\Pi$ , where  $\Pi$  is the matrix with elements  $\{\Pi_{n,m}\}$  in which the no null elements correspond to  $m = n, n - 1, n + 1$  and are given by Eqs. (3). In particular, using  $\frac{d\mathbf{P}}{dt} \approx \mathbf{P}(t + 1) - \mathbf{P}(t) = \mathbf{P}(t)(\Pi - \mathbf{1})$  the master equation can be obtained as

$$\begin{aligned} \frac{dP_n(t)}{dt} &= \Pi_{n-1,n} P_{n-1}(t) + \Pi_{n+1,n} P_{n+1}(t) \\ &\quad - (\Pi_{n,n} - 1) P_n(t), \end{aligned} \quad (4)$$

where the transition matrix is extended to include the special cases  $\Pi_{-1,0} = \Pi_{\kappa+1,\kappa} = 0$ . Expressing the time in units of  $\kappa$ —i.e.,  $\tilde{t} = t/\kappa$ —and rearranging terms, the master equation can be written as

$$\begin{aligned} \frac{dP_n(\tilde{t})}{d\tilde{t}} &= b^{(n-1)} P_{n-1}(\tilde{t}) + d^{(n+1)} P_{n+1}(\tilde{t}) \\ &\quad - (b^{(n)} + d^{(n)}) P_n(\tilde{t}), \end{aligned} \quad (5)$$

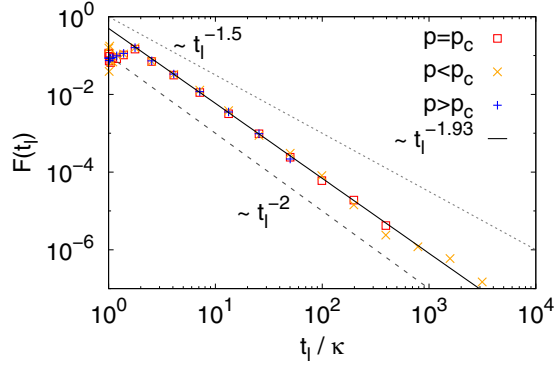


FIG. 2. Lifetime distribution calculated for the BMPG model using  $\kappa = 500$  for  $p = p_c = 1/\kappa$ ,  $p = 0.01 p_c$ , and  $p = 10 p_c$ . The lines correspond to power-laws with different exponents.

where  $b^{(-1)} = d^{(\kappa+1)} = 0$  and, for the other cases,

$$\begin{aligned} b^{(n)} &= n \left[ (1-p) - (1-p) \frac{n}{\kappa} \right] = n\beta(n) \\ d^{(n)} &= n \left[ 1 - (1-p) \frac{n}{\kappa} \right] = n[p + \beta(n)]. \end{aligned} \quad (6)$$

Here  $\beta(n) = (1-p)(1 - \frac{n}{\kappa})$ . When  $\beta$  is constant this master equation corresponds to a Galton-Watson process [32]. However, since the number of states is bounded in our model—i.e., the popularity inside the kernel of a given element cannot overstep the size of the kernel— $\beta$  is a decreasing function of  $n$ . In fitness landscape models, it is usually assumed that the total population size is fixed and determined by the carrying capacity of the environment [33]. Then, it is interesting to analyze the lifetime distribution for the elements in the kernel, since former studies used this kind of approach to model the lifetime distribution of species in ecologies [32]. As we will later show, different regimes arise in the BMPG model dynamics depending on the value of the product  $p\kappa$ . In Fig. 2 we show the lifetime distributions for the elements in the BMPG model for three different regimes:  $p > p_c$ ,  $p = p_c$ , and  $p < p_c$ , where  $p_c = 1/\kappa$ . These distributions, like all the results shown in this paper, are obtained by generating sequences of  $N = 10^7$  elements. In these series we compute the lifetime of each class of elements  $g$ , averaging over 20 realizations for each value of the parameters. In the three cases the lifetime of the elements is distributed according a long-tailed power-law distribution with exponent  $\approx -1.9$ . This exponent is between  $-3/2$ , which is the exit time of a random walk problem, and  $-2$ , the critical Galton-Watson branching process [32]. In particular, we expect an exponent different than  $-2$  because in the Galton-Watson process  $\beta(n)$  is constant and in our model a decreasing function of  $n$ . Furthermore, it has been suggested [34] that the lifetimes of taxa in the fossil record have a highly right-skewed distribution with a power-law tail with exponents in the range from  $-3/2$  to  $-2$ .

### C. Popularity distribution

As a further characterization of the model we calculate the popularity distribution  $P(s)$  of the generated sequences for the BMPG model using various values of the parameters  $p$

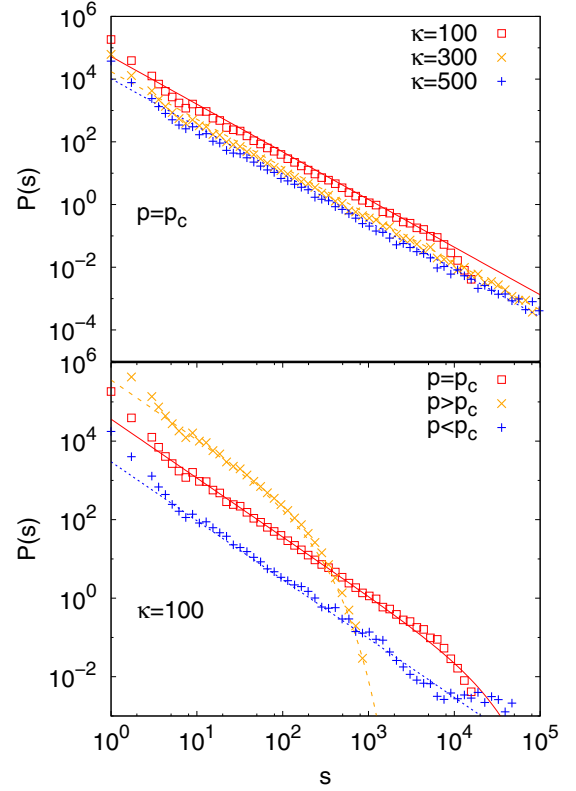


FIG. 3. Top panel: Popularity distribution calculated for the generated sequences with BMPG model for different kernel sizes. The straight lines correspond to fits of Eq. (7). In all cases the fitted exponents are  $\alpha \simeq 3/2$ . Bottom panel: Popularity distribution calculated for the generated sequences with the BMPG model for  $\kappa = 100$  and  $p = 0.1 p_c$ ,  $p_c$ , and  $10 p_c$ ; the full lines correspond to the estimation in Eq. (11).

and  $\kappa$ . In Fig. 3 (top panel) we show the results obtained for  $\kappa = 100, 300$ , and  $500$ , where in each case  $p = p_c = 1/\kappa$ . For all calculated  $P(s)$  we find that the distributions are very well fitted by a power law,

$$P(s) \sim s^{-\alpha}, \quad (7)$$

with  $\alpha \simeq 3/2$ , which means that the finite-sized kernel maintains the power-law decay found in the original CLS and Yule-Simon models, but it does modify the exponents  $\alpha$  in a particular form as we will discuss below.

To analyze the popularity distribution of elements in the BMPG model, we map the generation of elements of each class to a corresponding branching process. A branching process can be associated to a rooted tree. The process begins with a root node which creates children, and those children create children of their own and so on. The number of children that each node generates is represented by a stochastic variable. Moreover, each node belongs to a specific generation defined by its distance to the root node of the corresponding tree. In the scheme of the branching process, the generation of elements of a given class in the BMPG model is approximated by the evolution of a corresponding branching process. Each copy of an element of a given class corresponds to one of the nodes of a rooted tree, except for the root which corresponds to the first

apparition of this class of elements. In this way, each time the root is copied, a child of the first generation is born. In a similar manner, if a first generation node of this element is copied, a second generation child is born and so on.

Each element can have at most  $\kappa$  children, as it can only be copied while it is inside the memory kernel. In this process, the probability that a given node has  $i$  children can be well approximated by a binomial distribution,

$$p_i = \binom{\kappa}{i} \theta^i (1 - \theta)^{\kappa - i}, \quad (8)$$

where  $\theta = \frac{1}{\kappa}(1 - p)$  is the probability that a given element in the kernel is copied at every time step. In a branching process ruled by a binomial distribution, each link in the corresponding tree is generated with probability  $\theta$ . Since the number of links in a tree is  $s - 1$ , where  $s$  is the number of nodes or elements, then the probability of generating  $s - 1$  links is  $\theta^{s-1}$ . The number of missing links to complete a  $\kappa$ -ary tree with  $s$  internal nodes is  $\kappa s - (s - 1)$ , where internal means nodes that are not leaves. Then the probability that this process generates one particular  $\kappa$ -ary tree with  $s$  internal nodes and  $\kappa s - (s - 1)$  missing links is [35]  $\theta^{s-1} (1 - \theta)^{(\kappa - 1)s + 1}$ . The number of  $\kappa$ -ary trees with  $s$  internal nodes is [36]

$$N^{(\kappa)}(s) = \frac{1}{(\kappa - 1)s + 1} \binom{\kappa s}{s}, \quad (9)$$

and then the probability of having a  $\kappa$ -ary tree with  $s$  internal nodes and  $\kappa s - (s - 1)$  missing links, or in the context of the BMPG model an element of popularity  $s$ , is

$$P(s) = \frac{1}{(\kappa - 1)s + 1} \binom{\kappa s}{s} \theta^{s-1} (1 - \theta)^{(\kappa - 1)s + 1}. \quad (10)$$

Using Stirling's approximation for  $\kappa \gg 1$  and  $s \gg 1$  we can obtain the asymptotic behavior of Eq. (10),

$$P(s) \sim e^{-s/s_0(\kappa, \theta)} s^{-3/2}, \quad (11)$$

with  $s_0(\kappa, \theta) \approx \frac{\kappa - 1}{\kappa} \frac{2}{(1 - \kappa\theta)^2} = \frac{\kappa - 1}{\kappa} \frac{2}{p^2}$ .

In Fig. 3 (bottom panel) we show the resulting distributions  $P(s)$  calculated from the generated database with the BMPG model, together with the theoretical prediction Eq. (11), for the three different regimes. It can be seen that the theoretical prediction fits very well the distribution calculated from the sequences generated by the model, when the values of  $\kappa$  are large enough; this confirms that a branching process with a binomial distribution is a good approximation in this scenario. Moreover, when  $p = p_c = 1/\kappa$ , the exponential cutoff of distribution  $P(s)$  [Eq. (11)] has the form  $s_0 \propto \kappa^2$ . Therefore, for large values of  $\kappa$ , such as the ones used in Fig. 3 (top panel), the distribution tends to a power-law,  $P(s) \sim s^{-3/2}$ , in a wide range of values of  $s$  as observed in the figure.

Finally, we fitted the exponents of the frequency distribution for different values of the memory kernel extension  $\kappa$  and the probability of introducing a new element in the kernel (not shown) and we found that for small values of  $p$  all the exponents  $\alpha$  are nearly constant, while for larger values of  $p$  they grow as  $p$  increases. As expected, for  $p \leq 1/\kappa$ , all exponents have nearly the same value  $\alpha \approx 3/2$ .

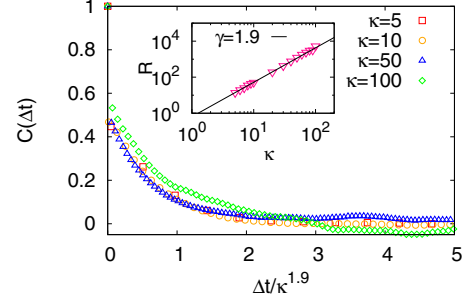


FIG. 4. Autocorrelation function  $C(\Delta t)$  calculated for sequences generated with the BMPG model for different kernel sizes and  $p = p_c$ , rescaled by  $\kappa^{1.9}$ . Inset: Decay characteristic time  $R$  of the autocorrelation function as a function of the memory kernel extension  $\kappa$ . The straight line corresponds to a linear function fit in log-log with exponent  $\gamma = 1.9$ .

### III. MEMORY EFFECTS

#### A. Correlations

As we mentioned in the Introduction, the sequences generated in the original Cattuto's model have long-range correlations. Since the BMPG model has a finite-sized memory kernel, we do not expect long-range correlations for the generated time series  $x[t]$ . To build the time series we assigned to each class of elements a random generated number from a Gaussian distribution [22]. To see how the correlation length behaves in this kernel, we computed the autocorrelation function  $C(\Delta t, t)$  of the time series  $x[t]$ ,

$$C(\Delta t, t) = \frac{\langle (x[t] - \mu(t))(x[t + \Delta t] - \mu(t + \Delta t)) \rangle}{\sigma(t)\sigma(t + \Delta t)}, \quad (12)$$

where  $\mu(t) = \langle x[t] \rangle$  and  $\sigma^2(t) = \langle (x[t] - \mu)^2 \rangle$  and  $\langle \dots \rangle$  means the expectation value. Considering the time series is stationary, then  $\mu(t) = \mu(t + \Delta t) = \mu_0$ ,  $\sigma(t) = \sigma(t + \Delta t) = \sigma_0$  and the autocorrelation function  $C(\Delta t, t) = C(\Delta t)$ , which can be computed as

$$C(\Delta t) = \frac{1}{N - \Delta t} \sum_{i=1}^{N - \Delta t} \hat{x}[t_i] \hat{x}[t_i + \Delta t], \quad (13)$$

where  $\hat{x}[t_i] = (x[t_i] - \mu_0)/\sigma_0$ ,  $\mu_0 = \frac{1}{N} \sum_i x[t_i]$  and  $\sigma_0^2 = \frac{1}{N} \sum_i (x[t_i] - \mu_0)^2$ . We found that  $C(\Delta t)$  decays exponentially,  $C(\Delta t) \sim e^{-\frac{\Delta t}{R}}$ , as can be seen in Fig. 4. We calculated the autocorrelation function for several values of  $\kappa$  and  $p = p_c = 1/\kappa$  and computed the correlation length  $R$  by fitting the exponential decay function. In the inset of Fig. 4 we show the correlation length  $R$  as a function of  $\kappa$ . As can be seen  $R$  grows as a power law with exponent  $\gamma = 1.9$ . Figure 4 also shows the collapse of the autocorrelation curves when the time is measured in units of the correlation length  $R \sim \kappa^{1.9}$ . Since the curves collapse as well if we use an exponent 2 instead of 1.9, we cannot ensure the precise value of this exponent. Finally, we also found that for any value  $p > p_c$  the correlation function  $C(\Delta t = 1)$  decays as  $\propto \kappa^{-1}$ . This result (not shown) means that the system decorrelates in this regime when  $\kappa \rightarrow \infty$ .



### B. Interevent time analysis

Interevent time analysis is common to many natural phenomena that exhibit memory such as earthquakes [2,37], sunspots [38], neuronal activity [39], and in several human activities [22,40,41]. In particular, following the analogy we already established between the occurrence of words in texts and our temporal series, the interevent time distribution can be analyzed in a similar manner than in texts [17]. All the elements in the sequence have a corresponding time of occurrence. Given an element of class  $g$  of the time series,  $t^{(g)}(j) \in \{1, 2, \dots, N\}$  corresponds to the time of the  $j$ th occurrence of  $g$ . Accordingly, the  $j$ th interevent time of a particular element  $g$  is defined as

$$\tau_j^{(g)} = t^{(g)}(j+1) - t^{(g)}(j), \quad (14)$$

If  $s^{(g)}$  is the total number of occurrences of  $g$  in the generated series, i.e., its popularity, then we can estimate the average inter-event time as  $\langle \tau^{(g)} \rangle \approx t_l^{(g)} / s^{(g)}$ , where  $t_l^{(g)}$  is the lifetime of the elements of class  $g$ , which is usually called the Zipf's wavelength of word  $g$  in text analysis [17]. For simplicity, let us write  $\tau$  instead of  $\tau^{(g)}$  and also introduce the probability density  $f(\tau)$  of interevent times for any particular class of elements in the series.

A bursty dynamics is characterized by a sequence of short periods of high activity followed by long periods of low activity [42]. The presence of burstiness in a time series can be studied through the probability distribution of interevent times  $f(\tau)$ , where any departure of  $f(\tau)$  from an exponential distribution may indicate the existence of memory in the interevent time process. For instance, in the case of words in a text a deviation is observed, which is usually well described by the stretched exponential distribution, or Weibull function [17],

$$f(\tau) = \frac{\beta}{\tau_0} \left( \frac{\tau}{\tau_0} \right)^{\beta-1} e^{-\left(\frac{\tau}{\tau_0}\right)^\beta}. \quad (15)$$

For this distribution  $\langle \tau \rangle = \tau_0 \Gamma(\frac{\beta+1}{\beta})$ , where  $\Gamma$  is the  $\Gamma$  function and  $0 < \beta \leq 1$ . If  $\beta = 1$ , burstiness is not present, as is the case for the generated series with CLS or the Yule-Simon's models [22]. On the other hand,  $0 < \beta < 1$  is an indicator of the presence of burstiness and, as the value of  $\beta$  approaches zero, the apparition of bursts in the time series increases. A more clear indication of burstiness comes when the distribution of interevent times is described by a power law,

$$f(\tau) = A\tau^{-\eta}, \quad (16)$$

where the degree of burstiness is characterized by the exponent  $\eta$  of the power law; the lower the exponent, the more bursty the dynamics is.

Finally, the presence of burstiness in empirical systems can also be characterized by calculating the coefficient of variation  $\sigma_\tau / \langle \tau \rangle$ , where  $\sigma_\tau$  is the standard deviation of the interevent times. Using this coefficient we compute the burstiness parameter  $B$  as [42]

$$B = \frac{(\sigma_\tau / \langle \tau \rangle) - 1}{(\sigma_\tau / \langle \tau \rangle) + 1} = \frac{\sigma_\tau - \langle \tau \rangle}{\sigma_\tau + \langle \tau \rangle}. \quad (17)$$

This parameter is greater than zero for a bursty dynamics and lower than zero when the dynamics becomes regular. When  $B = 0$  there is neither burstiness nor regularity like, for instance, in a Poisson process.

Using the approaches described above, we analyze the interevent time dynamics, i.e., burstiness, of several elements in series generated by the BMPG model. First, we define the level of activity of a given element  $g$  as the inverse of the Zipf's wavelength, as  $a_g = 1 / \langle \tau^{(g)} \rangle$ , where  $\langle \tau^{(g)} \rangle$  is the mean interevent time of this element. Then, analyzing the time series we found that, on average, the activity of the elements depends on their popularity through the following relation  $a \sim s^{1/2}$ , i.e., the most active elements are also the most popular. Second, we analyze the interevent time distributions for different sets of class of elements. Each set is built by aggregating the computed sequences of interevent times of each class of elements, in decreasing order of popularity, until reaching a specific number of total interevent times. In this way, we construct four sets of  $\sim 2 \times 10^6$  interevent times with decreasing level of activity.

In the top panel of Fig. 5 we show the distribution of interevent times  $f(\tau)$  for the most active set, using  $\kappa = 500$  and three different values of the parameter  $p$  ( $p = 0.1p_c$ ,  $p = p_c$ , and  $p = 10p_c$ ). As can be seen,  $f(\tau)$  has a power-law tail in the three cases with an exponent  $\eta = 3$ , which indicates a clear presence of burstiness. For  $p = p_c$  the whole distribution is well described by a power law. However, the deviation of  $p$  from  $p_c$  affects the distribution for short times. At  $p = 0.1p_c$  the series is homogeneous, i.e., composed almost entirely of elements of a single class, increasing the number of shorter interevent times. As  $p$  increases ( $p = 10p_c$ ), the number of different classes of elements also increases and the interevent times are longer, resulting in a flat distribution for lower values of  $\tau$ , as can be seen in this figure. We also observed that, independently of the value of  $\kappa$  used in the model (not shown), the decay exponent remains the same. For comparison, in this figure we also show the interevent time distribution obtained with Cattuto's model for  $\kappa_C = 500$  and  $p = p_c$  and the corresponding fit using a Weibull distribution with  $\beta = 1$ . As we mentioned, in this case the interevent time distribution is well explained by a Poisson process.

In the middle panel of Fig. 5 we show the distributions of interevent times  $f(\tau)$  for  $\kappa = 500$  and  $p = p_c$ , for several classes of elements grouped into subsets corresponding to different levels of average popularity. All distributions present a power-law decay with exponent  $\eta = 3$  and collapse when the axes are rescaled by the factor  $\langle \tau \rangle = \langle s \rangle^{-1/2}$ , where  $\langle s \rangle$  is the average of the popularity of the corresponding subset.

We also calculated the burstiness parameter  $B$  for different values of  $\kappa$  and  $p$ , using in this case the most active set—Fig. 5 (bottom panel). For all  $\kappa$  there is a range around  $\kappa p \sim 10$  where a bursty dynamics is evidenced, while for low values of  $\kappa p$ ,  $B < 0$ , since the series becomes more regular. As the distribution of interevent times for  $p = p_c$  is a power law with exponent  $\eta = 3$ , we can calculate the corresponding values of  $B$  for different values of  $\kappa$  by computing  $\langle \tau \rangle$  and  $\langle \tau^2 \rangle$  directly from the normalized distribution, resulting in

$$\langle \tau \rangle = \int_1^\kappa \tau f(\tau) d\tau = \frac{2\kappa}{\kappa + 1}, \quad (18)$$

and

$$\langle \tau^2 \rangle = \int_1^\kappa \tau^2 f(\tau) d\tau = \frac{2\kappa^2 \ln(\kappa)}{\kappa^2 - 1}. \quad (19)$$

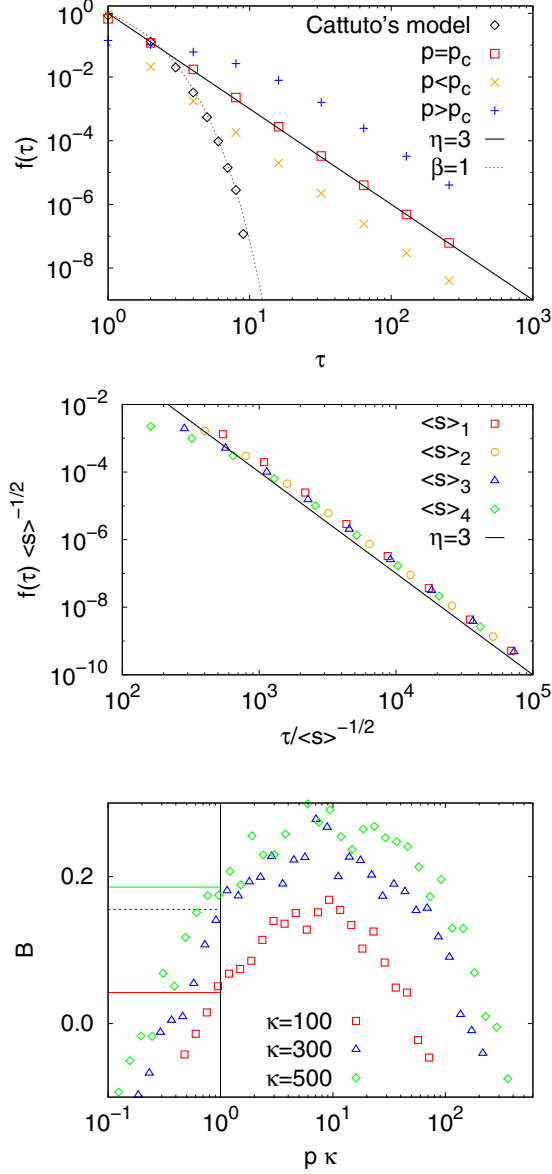


FIG. 5. Top panel: Distribution of interevent times for the BMPG model for with  $\kappa = 500$  and three values of  $p$ . We also show a fit according to Eq. (16) (black full line) for the case of  $p = p_c$ . For comparison, we show the distribution of interevent times for a sequence generated with CLS model for  $\kappa_c = 500$  and  $p = 1/\kappa_c$  and a fit with Eq. (15) (gray dashed line). Middle panel: Distribution of interevent times for different levels of activity, rescaled by  $\langle s \rangle^{-1/2}$  and a power law with exponent 3 for comparison. Bottom panel: Burstiness parameter  $B$  calculated for the BMPG model for various values of  $\kappa$  and  $p$  as a function of  $p\kappa$ ; the black vertical line corresponds to  $p\kappa = 1$  and the horizontal lines, to the values of  $B$  calculated from the distribution  $f(\tau)$ .

With these expressions we computed  $B$  [Eq. (17)] for different values of  $\kappa$ . The results are shown in the inset of Fig. 5 (bottom panel) as horizontal lines. As can be seen, the computed values of  $B$  coincide with the calculated values from the generated series with the BMPG model at  $p = p_c$ .

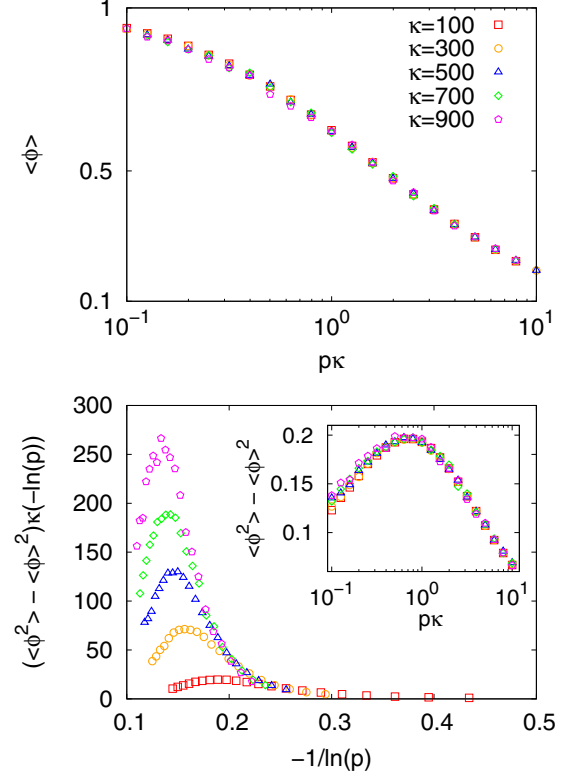


FIG. 6. Order parameter mean  $\langle \phi \rangle$  (top panel) and fluctuations  $(\langle \phi^2 \rangle - \langle \phi \rangle^2) \kappa [-\ln(p)]$  (bottom panel) calculated for the BMPG model for various values of  $\kappa$  as a function of  $p$  and rescaled by  $p_c = 1/\kappa$  (Inset).

#### IV. KERNEL ANALYSIS

To understand the emergence of burstiness in finite-sized kernels we analyzed the state inside the kernel in the BMPG model as a function of the model's parameters. In analogy with phase transitions, in a first approach we analyzed the state inside the kernel by defining an order parameter  $0 \leq \phi \leq 1$  as follows:

$$\phi = \frac{1}{\kappa} \max_g n_{\kappa}^{(g)}, \quad (20)$$

where  $n_{\kappa}^{(g)}$  is the popularity of the class of elements  $g$  inside the kernel of length  $\kappa$ . According to this definition, when the kernel is full of elements of a single class, the order parameter is equal to one. On the other hand, when all elements in the kernel are of different classes  $\phi = 1/\kappa$ , taking its minimum value. In Fig. 6 (top panel) we show the mean value of the order parameter as a function of  $p\kappa$  for several values of the kernel extension  $\kappa$ , where  $\phi$  is computed inside  $N/\kappa$  disjoint segments of length  $\kappa$  all across the generated series ( $N$  being the length of the series), and the mean  $\langle \phi \rangle$  is calculated over all  $N/\kappa$  segments. One can note that all the curves collapse when plotted as a function of  $\kappa p$ , also it can be observed a continuous transition from a disordered state  $\phi \sim 1/\kappa \approx 0$  to an ordered state  $\phi \approx 1$  as  $\kappa p$  decreases. The inflexion of the curves that occurs at  $p\kappa \approx 1$  suggests a transition at this point.

To test the existence of some form of criticality in the transition we estimate the fluctuations of the order parameter

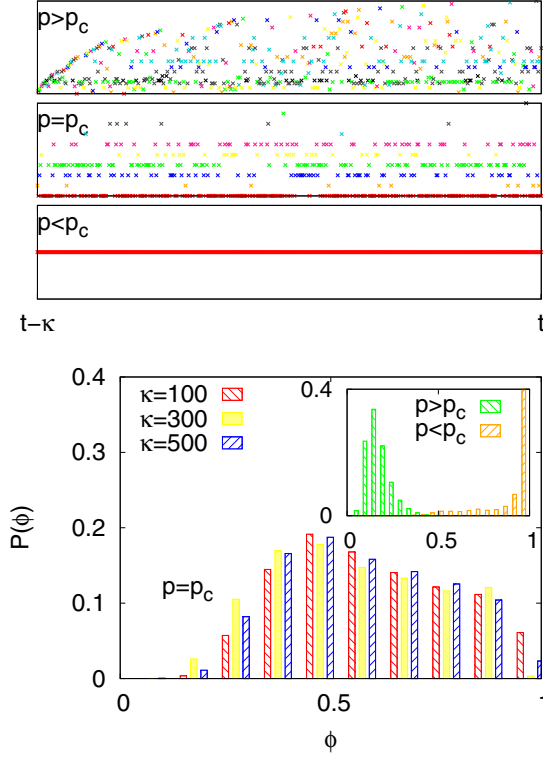


FIG. 7. Top panel: kernel configurations for the supercritical ( $p > p_c$ ), critical ( $p = p_c$ ), and subcritical ( $p < p_c$ ) states. Bottom panel: Order parameter distribution calculated for the BMPG model for various values of  $\kappa$  in the case of  $p = p_c$ . Inset: Examples of the order parameter distribution in the case of  $p > p_c$  and  $p < p_c$ .

by computing the variance  $\sigma_\phi^2 = \langle \phi^2 \rangle - \langle \phi \rangle^2$ . The variance as a function of  $p\kappa$  shows a peak at  $p\kappa \approx 1$ —see inset of the bottom panel of Fig. 6—indicating that fluctuations reach a maximum at the inflexion point of the order parameter. Since all the curves collapse when plotted as a function of  $p\kappa$ , the magnitude of the fluctuations is independent of the size of the kernel. Using a heuristic argument we can make an analogy with statistical mechanics by associating  $p$  to the temperature—since this variable introduces disorder in the kernel—and the kernel extension  $\kappa$  to the size of the system. If we think that the probability of introducing a new element results from an activated process, then  $p \propto \exp(-\Delta E/T)$ , where  $\Delta E$  is the activation energy. From this we obtain  $T \propto -1/\ln(p)$ , or in a simple assumption  $T = -1/\ln(p)$ , which introduces all the temperature range. In this scheme we can also define the susceptibility as  $\chi = \frac{\kappa}{T} \sigma_\phi^2 = -\ln(p)\kappa \sigma_\phi^2$ . A plot of the susceptibility  $\chi$  as a function of  $T = -1/\ln(p)$  is shown in Fig. 6 (bottom panel), where it can be seen that the peak of the susceptibility grows with the size of the kernel and is expected to diverge at  $T = 0$  in the thermodynamic limit  $\kappa \rightarrow \infty$ , resembling the behavior of, for instance, the one-dimensional Ising model.

In the top panel of Fig. 7 we show kernel configurations for the three main states: supercritical ( $p > p_c$ ), critical ( $p = p_c$ ), and subcritical ( $p < p_c$ ). In the subcritical state, the kernel is full with elements of the same class ( $\phi \sim 1$ ); in the supercritical state, there are many elements coexisting in the kernel with

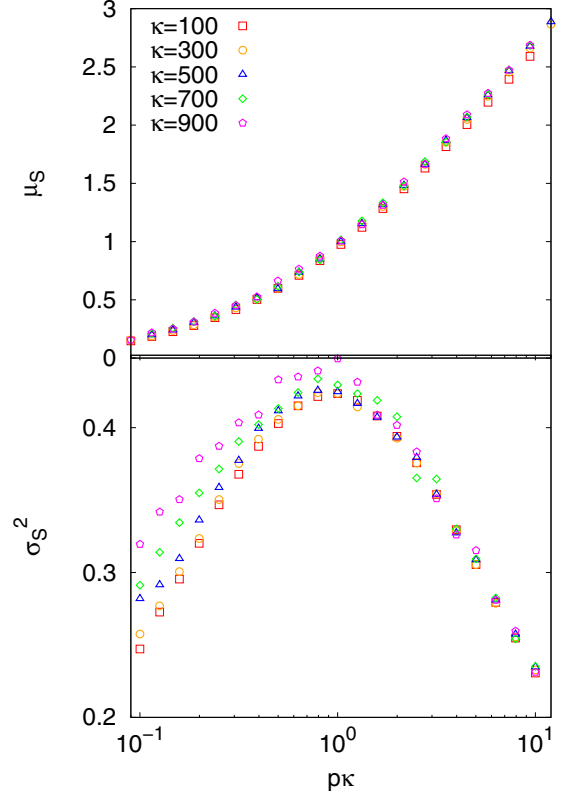


FIG. 8. Entropy mean  $\mu_S$  (top panel) and fluctuations  $\sigma_S$  (bottom panel) calculated for the BMPG model for various values of  $\kappa$  as a function of  $p\kappa$ .

similar weights and in the critical case one can distinguish the most popular element by inspection. Finally, a better description of the fluctuations of  $\phi$  in the kernel can be obtained by calculating the distribution of the order parameter. We computed the distributions at the point where the fluctuations are maximum, i.e.,  $p = p_c = 1/\kappa$ , for several values of  $\kappa$ . In the bottom panel of Fig. 7 we show these distributions corresponding to  $\kappa = 100, 300$ , and  $500$ . The distributions do not depend on the value of  $\kappa$  and broaden around  $\phi = 0.5$ , further confirming the results obtained when measuring the variance. In the inset we show distributions for the subcritical case  $p < p_c$  and the supercritical case  $p > p_c$  for the case of  $\kappa = 500$  with  $p = 0.1p_c$  and  $p = 10p_c$ . Both distributions are narrow, the first one is peaked close to 1 and the second one close to 0, as expected.

The distribution of elements inside the kernel can be studied by computing the kernel entropy

$$S = - \sum_g \frac{n_\kappa^{(g)}}{\kappa} \ln \left( \frac{n_\kappa^{(g)}}{\kappa} \right), \quad (21)$$

where, as before,  $n_\kappa^{(g)}$  is the popularity of the  $g$ th class of elements inside the kernel. In Fig. 8 we show the mean value of the entropy,  $\mu_S$  (top panel) and its fluctuations,  $\sigma_S$  (bottom panel), as a function of  $p\kappa$ , and for several values of  $\kappa$ . Similar to what was observed for the order parameter, all the entropy curves collapse when plotted as a function of  $p\kappa$ . The variance of the entropy behaves as the fluctuations of the

order parameter, reaching a maximum value at  $p \sim p_c$ , also indicating the existence of a transition. However, it seems to increase with the size of the kernel. The entropy fluctuations are related to the presence of burstiness, since they imply that the rate at which an element is copied is a varying quantity.

It is interesting to note that the bursty activity of solar flares has a distribution of interevent times with a power-law tail with exponent  $\sim 3$ , in accordance with the results of BMPG model. The distribution of solar flares has been satisfactorily explained using a time-dependent Poisson process that results from the superposition of piecewise-constant Poisson processes [43]. Within this approach, the process is decomposed in time intervals, in which the interevent time is consistent with a constant rate Poisson process. In analogy with solar flares, the copy rate of a given class of elements varies in the BMPG model when it is near the transition, as evidenced by the fluctuations of the entropy and the order parameter. The mentioned mechanism is robust against variations of the distribution of Poisson rates. Hence, it can be easily adapted to explain the distribution of interevent times of the BMGP model.

## V. DISCUSSION

In this work we have studied and characterized a preferential growth stochastic model with a bounded memory kernel. Specifically, we have modified the Yule-Simon stochastic process by introducing a finite-sized memory kernel of extension  $\kappa$  and called this model the BMPG model. We studied several statistical properties of the series of elements generated by BMPG model using numerical simulations and standard tools from stochastic processes.

We found that the lifetime distributions of the different classes of elements in the series follow a power law. We derived the master equation that rules the probability of having  $n$  copies of a given class in the kernel at time  $t$  and we found that this equation is similar to the ones proposed for models of species lifetime in ecologies [32], which consists in a birth and death process. Moreover, the exponents of the obtained distributions from the BMPG model,  $\sim 1.9$ , are in the range of the reported in empirical systems [34].

The BMPG model also generates elements whose popularity is distributed according to a power law. This is in accordance with the highly skewed and long-tailed distributions found in the original Yule-Simon model and in the modified version introduced by Cattuto, Loreto, and Servedio, which includes a long-tailed memory kernel. In particular, the exponent of the distribution ( $3/2$ ) in the case of the finite-sized kernel can be explained in terms of a branching process ruled by a binomial distribution, where the maximum number of children each node can have is equal to the size of the kernel. This approach works well for large enough kernels and is also able to explain the exponential cutoff observed in the distributions obtained with our model.

As expected, the correlations observed in the series of elements produced by BMPG model are short-ranged, unlike the long-range correlations observed in CLS model. However, the correlation length in the finite-sized kernel grows nearly quadratically with the kernel extension.

An interesting effect observed in the presence of a finite-sized kernel—which is not observed neither in the Yule-Simon

model nor in CLS model—is that the tail of the interevent time distribution of elements of the series, with a defined level of activity, decays as a power law with exponent  $\sim 3$  and is independent of the activity level of the pool of elements. This means that the generated sequences of a pool of elements show bursts of activity followed by periods of latency, at least for a given range of the model parameters. Moreover, the interevent time distributions collapse when rescaled by the level of activity of the pool of elements and, in this respect, we also found that the level of activity increases with the popularity of the elements in the series. The burstiness parameter is greater than zero in a range of the model's parameters, reinforcing the evidence of the presence of burstiness in the generated series. The mechanism of this bursty dynamics can be associated to a superposition of Poisson processes with a particular distribution of rates; in particular, this mechanism was used to explain the interevent time distribution of solar flares [43] that, like in BMPG model, is distributed by a power law with exponent 3. In contrast, our previous studies of the Yule-Simon and CLS models showed that the interevent time distribution is well explained by a Poisson process with a single rate [22].

To explain the presence of the bursty dynamics in the sequences generated by the BMPG model, we characterized the state of the memory kernel defining an order parameter that measures the occupation ratio of the kernel by the most popular class of elements. Also, as a measure of the distribution of classes of elements inside the kernel, we computed the kernel entropy. Studying the mean value and fluctuations of the order parameter and the entropy, we found that the state of the kernel goes through a transition from an ordered state (low values of  $p$ ) to a disordered state (high values of  $p$ ) and that there is a critical point at  $p = p_c = 1/\kappa$ , where the fluctuation of both quantities are maximum. Particularly, entropy fluctuations can be closely related to the emergence of burstiness in BMPG model, since it implies that the rate at which an element is copied is a varying quantity. Moreover, since all the curves collapse when plotted as a function of  $p\kappa$ , the magnitude of the fluctuations of the order parameter and the entropy result independent of the size of the kernel.

In principle, the Yule-Simon model seems to be a limiting case of the BMPG when  $\kappa \rightarrow \infty$ . However, this is not the case because at large times both models have different dynamics. This fact explains the existence of the memory effects observed in BMPG, which are absent in Yule-Simon. More precisely, we study temporal series generated by BMPG once the kernel is in a stationary state, i.e., after a large enough transient time  $t$  ( $t \gg \kappa$ ). In contrast, the temporal series generated by the Yule-Simon model are in a nonstationary state because the kernel is continually growing. As a consequence, the time series generated by the Yule-Simon model at time  $t = \kappa$  result different from the series generated by our model with a kernel of size  $\kappa$ . In particular, the correlations lengths, which we found to grow as  $\sim \kappa^2$ , is expected to diverge in the Yule-Simon model; however, these correlations are studied at  $p = p_c$  where  $p_c = 1/\kappa$ . In the Yule-Simon model,  $p_c \rightarrow 0$  in the limit of  $\kappa \rightarrow \infty$ , meaning that the Yule-Simon model is in a disordered state for any finite  $p > 0$  and sufficiently long times  $t$ , but we found that the correlation vanishes in the disordered state. This explains why we do not observe a diverging correlation length in Yule-Simon process. It also explains the lack of burstiness



in this model since, as it was shown, a bursty dynamics is only observed near the critical point.

Finally, in complementary results not shown in this paper, in which we repeated all the analysis using an exponential short-ranged memory kernel  $Q(\Delta t) = \frac{1}{\kappa_e} \exp(-\Delta t/\kappa_e)$  with decaying rate  $\kappa_e$ , we found similar behaviors for the popularity and interevent time distributions to those obtained with the finite-sized kernel. This suggests that these statistical properties are independent of the specific functional form of the

memory kernel, depending instead on how fast the distribution decays, i.e., on the asymptotic behavior for large  $\Delta t$ .

## ACKNOWLEDGMENT

This work was partially supported by grants from CON-ICET (PIP Grant No. 112 20150 10028), and SeCyT Grant No. 05/B370—Universidad Nacional de Córdoba (Argentina).

- 
- [1] M. E. J. Newman, *Contemp. Phys.* **46**, 323 (2005).
  - [2] P. Bak, K. Christensen, L. Danon, and T. Scanlon, *Phys. Rev. Lett.* **88**, 178501 (2002).
  - [3] E. T. Lu and R. J. Hamilton, *Astrophys. J.* **380**, L89 (1991).
  - [4] G. K. Zipf, *J. Clin. Psychol.* **6**, 306 (1950).
  - [5] J. C. Willis and G. U. Yule, *Nature* **109**, 177 (1922).
  - [6] V. Pareto, *Cours d'économie politique* (Librairie Droz, Genève, 1964), Vol. 1.
  - [7] B. Blasius and R. Tönjes, *Phys. Rev. Lett.* **103**, 218701 (2009).
  - [8] M. Mitzenmacher, *Internet Math.* **1**, 226 (2003).
  - [9] B. Corominas-Murtra, R. Hanel, and S. Thurner, *Proc. Natl. Acad. Sci. USA* **112**, 5348 (2015).
  - [10] G. U. Yule, *Philos. Trans. R. Soc. London B* **213**, 21 (1925).
  - [11] H. A. Simon, *Biometrika*, 425 (1955).
  - [12] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
  - [13] V. N. Livina, S. Havlin, and A. Bunde, *Phys. Rev. Lett.* **95**, 208501 (2005).
  - [14] E. G. Altmann and G. Cristadoro, *Proc. Natl. Acad. Sci. USA* **109**, 11582 (2012).
  - [15] A. L. Schaigorodsky, J. I. Perotti, and O. V. Billoni, *Physica A* **394**, 304 (2014).
  - [16] M. A. Montemurro and P. A. Pury, *Fractals* **10**, 451 (2002).
  - [17] E. G. Altmann, J. B. Pierrehumbert, and A. E. Motter, *PLoS ONE* **4**, e7678 (2009).
  - [18] K. Tanaka-Ishii and A. Bunde, *PLoS ONE* **11**, 1 (2016).
  - [19] J. I. Perotti, H. Jo, A. L. Schaigorodsky, and O. V. Billoni, *Europhys. Lett.* **104**, 48005 (2013).
  - [20] C. Cattuto, V. Loreto, and V. D. P. Servedio, *Europhys. Lett.* **76**, 208 (2006).
  - [21] C. Cattuto, V. Loreto, and L. Pietronero, *Proc. Natl. Acad. Sci. USA* **104**, 1461 (2007).
  - [22] A. L. Schaigorodsky, J. I. Perotti, and O. V. Billoni, *PLoS ONE* **11**, 1 (2016).
  - [23] M. Simkin and V. Roychowdhury, *Phys. Rep.* **502**, 1 (2011).
  - [24] D. Wang, C. Song, and A.-L. Barabási, *Science* **342**, 127 (2013).
  - [25] P. Lansky, F. Polito, and L. Sacerdote, *J. Stat. Phys.* **165**, 661 (2016).
  - [26] A. Garavaglia, R. van der Hofstad, and G. Woeginger, *J. Stat. Phys.* **168**, 1137 (2017).
  - [27] P. L. Krapivsky and S. Redner, *Phys. Rev. E* **63**, 066123 (2001).
  - [28] G. Bianconi and A.-L. Barabási, *Europhys. Lett.* **54**, 436 (2001).
  - [29] S. Lehmann, A. D. Jackson, and B. Lautrup, *Europhys. Lett.* **69**, 298 (2005).
  - [30] R. Lambiotte, *J. Stat. Mech.* (2007) P02020.
  - [31] M. Medo, G. Cimini, and S. Gualdi, *Phys. Rev. Lett.* **107**, 238701 (2011).
  - [32] S. Pigolotti, A. Flammini, M. Marsili, and A. Maritan, *Proc. Natl. Acad. Sci. USA* **102**, 15747 (2005).
  - [33] B. Drossel, *Adv. Phys.* **50**, 209 (2001).
  - [34] K. Sneppen, P. Bak, H. Flyvbjerg, and M. H. Jensen, *Proc. Natl. Acad. Sci. USA* **92**, 5209 (1995).
  - [35] F. F.-C. Á. Corral, in *Self-Organized Criticality Systems*, edited by M. J. Aschwanden (Open Academic Press, Berlin, 2014), Chap. 5.
  - [36] M. Drmota, *Cubo J.* **6**, 2004 (2004).
  - [37] A. Corral, *Phys. Rev. Lett.* **92**, 108501 (2004).
  - [38] M. S. Wheatland, P. A. Sturrock, and J. M. McTiernan, *Astrophys. J.* **509**, 448 (1998).
  - [39] J. Keat, P. Reinagel, R. Reid, and M. Meister, *Neuron* **30**, 803 (2001).
  - [40] A.-L. Barabási, *Nature* **435**, 207 (2005).
  - [41] H.-H. Jo, J. I. Perotti, K. Kaski, and J. Kertész, *Phys. Rev. E* **92**, 022814 (2015).
  - [42] K.-I. Goh and A.-L. Barabási, *Europhys. Lett.* **81**, 48002 (2008).
  - [43] M. S. Wheatland, *Astrophys. J. Lett.* **536**, L109 (2000).