Finite-size scaling of survival probability in branching processes

Rosalba Garcia-Millan,[1,2] Francesc Font-Clos,[1,3] and Álvaro Corral[1,3]
[1]*Departament de Matemàtiques, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain*
[2]*Departament de Física, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain*
[3]*Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain*
(Received 14 November 2014; published 20 April 2015)

Branching processes pervade many models in statistical physics. We investigate the survival probability of a Galton-Watson branching process after a finite number of generations. We derive analytically the existence of finite-size scaling for the survival probability as a function of the control parameter and the maximum number of generations, obtaining the critical exponents as well as the exact scaling function, which is $\mathcal{G}(y) = 2ye^y/(e^y - 1)$, with y the rescaled distance to the critical point. Our findings are valid for any branching process of the Galton-Watson type, independently of the distribution of the number of offspring, provided its variance is finite. This proves the universal behavior of the finite-size effects in branching processes, including the universality of the metric factors. The direct relation to mean-field percolation is also discussed.

DOI: 10.1103/PhysRevE.91.042122 PACS number(s): 05.70.Jk, 45.70.Ht, 05.65.+b

I. INTRODUCTION

Branching processes have become a very useful modeling tool [1,2], originally in demography and population biology [3], later in genetics and in the theory of nuclear reactions [4,5], and more recently in seismology [6,7]. Important applications in statistical physics have arisen due to the relationships of branching processes with critical phenomena, through percolation theory and self-organized criticality (SOC) [8–12].

Indeed, branching processes provide one of the simplest examples of a second-order phase transition, equivalent to percolation in the Bethe lattice, and therefore free of geometric complexity [10]. These phase transitions (also called *continuous*) are characterized by a sudden change of an "order parameter" from zero to a nonzero value at precisely a critical value of a "control parameter" [13,14]. This fact is exploded in SOC theory, with mean-field sand-pile models where avalanches propagate through a system by means of a "critical branching process"; the peculiarity of SOC is that the critical state is reached in a spontaneous, self-organized way [10,15–17].

The common fact of criticality is the fulfillment of scaling laws for the thermodynamic variables and the correlation length ξ close to the critical point; e.g., in a magnetic system [13],

$$\frac{m}{|t|^\beta} = g_\pm\left(\frac{h}{|t|^\Delta}\right),$$

and an analogous one for $\xi|t|^\nu$, where m is the dimensionless magnetic moment per particle, which plays the role of order parameter; t is the reduced temperature; h is the reduced magnetic field; β, Δ, and ν are critical exponents; and $g_\pm$ represents two scaling functions, one $(+)$ for $t > 0$ and another one $(-)$ for $t < 0$. The critical point is achieved at $t = h = 0$.

A fundamental approach to analyze critical phenomena is by means of finite-size scaling. It turns out to be that the sharp change in the properties of a system at a critical state is only possible in the thermodynamic limit (this is the limit of infinite system size). However, in practice, computer simulations cannot attain such a limit, for obvious reasons, and one cannot infer the existence of a critical point from the results of computer simulations alone. Then, in a finite system of size L an additional dependence appears, given by the ratio ξ/L, which can be replaced by $L|t|^\nu$, yielding the ansatz

$$m = |t|^\beta \tilde{g}_\pm\left(\frac{h}{|t|^\Delta}, L|t|^\nu\right) = \frac{1}{L^{\beta/\nu}}\mathcal{G}(aL^{\Delta/\nu}h, bL^{1/\nu}t), \quad (1)$$

where $\tilde{g}_\pm$ and $\mathcal{G}$ are bivariate scaling functions and a and b are metric factors introduced to ensure universality [18]. The previous scaling law is known as finite-size scaling. Note that, in the case when t and h appear linearly as arguments of the scaling function, no distinction is made between $t > 0$ and $t < 0$ and a single scaling function $\mathcal{G}$ is enough for describing both regimes. The reason is that in a finite system there is no singularity at $t = h = 0$, where $\mathcal{G}$ is smooth and analytic [18].

In this paper we show that branching processes with size limitations display finite-size scaling in the same way as in critical phenomena. We are able to derive the exact form of the scaling function as well as the critical exponents. In the next section we review the basic language for branching processes, whereas in the third one we apply it to finite-size branching processes. In the last section some implications are discussed.

Usually, when the distance to the critical point is kept fixed, the decay of the probability of surviving towards the infinite-size case is exponential in L [1,2]. In contrast, the finite-size scaling approach demonstrates, keeping fixed the distance to the critical point in relative units of $1/L$, a power-law decay with L, resulting in a sort of law of corresponding states which turns out to be valid for any system size L (provided that this is large enough) and for any branching process of the Galton-Watson type (provided a finite variance). This universality [14,19] arises because, as we show, at the critical point the only relevant quantity is the variance of the distribution that defines the Galton-Watson process.

II. OVERVIEW OF PREVIOUS IMPORTANT RESULTS

For the connection of branching processes with critical phenomena it is enough to consider simply the Galton-Watson process [1]. This is started, in the zeroth generation of the process, by one single element, which produces other elements, called *offspring*, in a number given by a random variable K. The offspring of the initial element constitute the first generation of the process, and each one of them produces again a random number of offspring, which are the second generation, and so on. The main ingredient of the model is that the number of offspring of any element follows the same distribution (that of K), and each of these random numbers is independent from those of the other elements.

The first variable of interest is N_t, which represents the number of elements in each generation t. The initial condition is written then as $N_0 = 1$. The key question in branching processes is if the process gets extinct or not, and this is represented by the event $N_t = 0$; in particular, as extinction is an absorbing state, all extinction events are included in $N_t = 0$ with $t \to \infty$. Then the probability of extinction is

$$P_{\text{ext}} = \lim_{t\to\infty} \text{Prob}(N_t = 0).$$

At this point it is very useful to introduce the probability generating function. Consider a generic discrete random variable X which takes value 0 with probability p_0, value 1 with p_1, and value x with probability p_x. The probability generating function $f_X(z)$ is defined as

$$f_X(z) = \sum_{x=0}^{\infty} p_x z^x = \langle z^X \rangle,$$

where the dependence of $f_X(z)$ is on z, and X indicates to which random variable it corresponds (in our case $X = K$ or $X = N_t$). Useful but straightforward properties of $f_X(z)$ are the following: (i) $f_X(0) = p_0$, (ii) $f_X(1) = 1$, (iii) $f'_X(1) = \langle X \rangle$, (iv) $f''_X(1) = \langle X(X-1) \rangle$, (v) $f'_X(z) \geqslant 0$ in [0,1], and (vi) $f''_X(z) \geqslant 0$ in [0,1], where the primes denote differentiation (with respect to the variable z), and we assume those expected values exist and are finite.

Applying the first of these properties to the variable N_t, we get, for the probability of extinction,

$$P_{\text{ext}} = \lim_{t\to\infty} f_{N_t}(0),$$

which, as we will see, constitutes a great simplification of the calculation. Using that $N_{t+1} = \sum_{i=1}^{N_t} K_{t,i}$, where $K_{t,i}$ is the number of offspring of the ith element in the tth generation, and that $f_{N_1}(z) = f_K(z)$, it is possible to derive a fundamental theorem in branching processes, which is

$$f_{N_t}(z) = f_K(f_K(\dots f_K(z)\dots)) = f_K^t(z),$$

where the superindex t means composition t times (not power). This is valid because the $K_{t,i}$ are independent and identically distributed [1,17]. Therefore, the probability of extinction turns out to be

$$P_{\text{ext}} = \lim_{t\to\infty} f_K^t(0),$$

i.e., the repeated iteration of the origin $z = 0$ through the function $f_K(z)$.

Using the rest of properties of probability generating functions listed above it is possible to show that the probability of extinction is given by $P_{\text{ext}} = q$, where q is the smallest non-negative fixed point of $f(z)$, i.e., we have $f_K(q) = q$. For $\langle K \rangle \leqslant 1$ it turns out to be that $q = 1$ but for $\langle K \rangle > 1$ one gets $0 \leqslant q < 1$ [1,2,17]. As q varies continuously with $\langle K \rangle$ this reveals the existence of a continuous (or second-order) phase transition in the system, with control parameter $\langle K \rangle$ and with the probability of surviving to extinction, or survival probability, $P_{\text{surv}} = 1 - P_{\text{ext}} = 1 - q$, behaving as an order parameter. This is zero below and at the critical point $\langle K \rangle = 1$, and strictly positive above the critical point, when q and therefore P_{surv} will depend on the parameters of the distribution of K. The phase diagram of the Galton-Watson process consists then of three regimes: subcritical, critical, and supercritical, depending on the value of $\langle K \rangle$.

III. FINITE-SIZE EFFECTS

In contrast with that explained above, we consider here a system with a limitation in size, that is, the limit of infinite generations $t \to \infty$ cannot be reached and one has instead an imposed maximum number of generations L, which we identify with system size. We identify therefore $t = L$ with the boundary of the system. So all extinction events are included in the boundary-extinction event, and then the probability of extinction is given by $P_{\text{ext}}(L) = \text{Prob}(N_L = 0)$. One can see that this value will be smaller than in an infinite equivalent system, as extinction at $t = L$ is a particular case of extinction at $t \to \infty$. In any case, as in the infinite system, we will have that the probability is given by the iteration of the origin through $f(z) = f_K(z)$ [from now on, to ease the notation, we drop the subindex K in $f_K(z)$], i.e.,

$$P_{\text{ext}}(L) = f_{N_L}(0) = f^L(0).$$

So the fixed point q will not be reached but instead we will have that $P_{\text{ext}}(L) < q$.

What we expect is that after a large number of generations n (if L is also large), $f^n(0)$ will be close to the fixed point q, i.e., $q - f^n(0)$ will be close to zero. Then we will perform a Taylor expansion of $f[f^n(0)]$, considered a function of $f^n(0)$, around the abscissa point q, that is,

$$\begin{aligned} f^{n+1}(0) &= f[f^n(0)] \\ &= f(q) + f'(q)[f^n(0) - q] \\ &\quad + \tfrac{1}{2} f''(q)[f^n(0) - q]^2 + \dots, \end{aligned}$$

up to second order. Using the fixed-point condition we can rewrite

$$q - f^{n+1}(0) = f'(q)[q - f^n(0)] - \tfrac{1}{2} f''(q)[q - f^n(0)]^2;$$

in other words, the distance d to the fixed point q at iteration $n + 1$ is given by

$$d_{n+1} = M d_n - C d_n^2$$

when n is large, defining $M = f'(q)$ and $C = f''(q)/2$. Actually, it is simpler to iterate the inverse of the distance, which at the lowest orders in d_n is

$$\frac{1}{d_{n+1}} = \frac{1}{M d_n}\left(1 + \frac{C d_n}{M}\right),$$

and introducing the inverse of the distance, $c_n = 1/d_n$,

$$c_{n+1} = \frac{c_n}{M} + \frac{C}{M^2}.$$

It is easy to see that successive iterations lead to

$$c_{n+\ell} = \frac{c_n}{M^\ell} + \frac{C}{M^2}\left(1 + \frac{1}{M} + \cdots + \frac{1}{M^{\ell-1}}\right)$$
$$= \frac{c_n}{M^\ell} + \frac{C(1-M^\ell)}{M^{\ell+1}(1-M)}, \quad (2)$$

using the formula of the geometric progression. From here one can see that, for fixed M and large ℓ, the decay of the distance to the fixed point q is proportional to M^ℓ, and therefore exponential in ℓ (we will see below that $M < 1$ except at the critical point).

A. Subcritical and critical cases

Now we need to consider separately the three regimes. First we deal with the subcritical phase, defined by $\langle K\rangle < 1$, for which the fixed point is $q = 1$ and then, by the properties of the probability generating function, $M = f'(1) = \langle K\rangle$ and $2C = f''(1) = \langle K(K-1)\rangle = \sigma^2 + \langle K\rangle(\langle K\rangle - 1)$, where σ^2 is the variance of K. This leads to

$$c_{n+\ell} = \frac{c_n}{\langle K\rangle^\ell} + \frac{\sigma^2(1-\langle K\rangle^\ell)}{2\langle K\rangle^{\ell+1}(1-\langle K\rangle)} - \frac{1-\langle K\rangle^\ell}{2\langle K\rangle^\ell},$$

where we will introduce the rescaled distance to the critical point (distance in units of $1/\ell^{1/\nu}$), as

$$y = \ell^{1/\nu}(\langle K\rangle - 1),$$

this means that $\langle K\rangle = 1 + y/\ell^{1/\nu}$. Further, we will take $\ell \to \infty$, and then an interesting limit emerges for $\nu = 1$, yielding $\langle K\rangle^\ell \to e^y$. In order to keep y finite we will impose that we are in the vicinity of the critical point, $\langle K\rangle \to 1^-$. Notice that in this limit only the middle term in the expression for $c_{n+\ell}$ survives, i.e.,

$$c_{n+\ell} = -\frac{\sigma^2(1-e^y)\ell}{2e^y y},$$

and from here we can obtain the probability of surviving in a system of limited size $L = n + \ell$ as

$$P_{\rm surv}(L) = 1 - P_{\rm ext}(L) = 1 - f^L(0) = \frac{1}{c_L} = \frac{2e^y y}{\sigma_c^2(e^y-1)L},$$

where we have considered $L \gg n$, and so $\ell \simeq L$, and also have taken the variance right at the critical point, σ_c^2.

One can realize that this result also includes the critical case, given by $M = \langle K\rangle = 1$, just taking the limit $y \to 0$, for which $2ye^y/(e^y-1) \to 2$. This is in correspondence with the replacement of Eq. (2) by $c_{n+\ell} = c_n + C\ell$. Therefore,

$$P_{\rm surv}(L) = \frac{2}{\sigma_c^2 L},$$

at the critical point. This was apparently first proved by Kolmogorov under more restrictive assumptions [1, p. 20], [2, p. 19], [4, p. 47].

B. Supercritical case

We show now that the supercritical case leads, through a more involving path, to exactly the same result as the subcritical case. The main difference is that M is no longer the derivative of $f(z)$ at $z = 1$ but at $z = q < 1$. Let us expand by Taylor $f(q)$ around $q = 1$, i.e., close to the value of the fixed point corresponding to the critical point,

$$f(q) = 1 + \langle K\rangle(q-1) + \frac{f''(1)}{2}(q-1)^2 = q,$$

which leads, for $q \neq 1$, to

$$q = 1 - \frac{2}{f''(1)}(\langle K\rangle - 1). \quad (3)$$

Substituting in the Taylor expansion of $f'(q)$, up to first order in $q - 1$,

$$M = f'(q) = \langle K\rangle + f''(1)(q-1) = 2 - \langle K\rangle,$$

which is smaller than 1 as $\langle K\rangle > 1$, and using that $\langle K\rangle = 1 + y/\ell$, allows one to calculate $M^\ell = e^{-y}$ and $1 - M = y/\ell$. Notice that this looks the same as in the subcritical case but replacing y by $-y$. Going back to Eq. (2),

$$c_{n+\ell} = \frac{\sigma^2(1-e^{-y})\ell}{2e^{-y}y},$$

when ℓ is very large, approximating $2C = f''(q) = f''(1) = \sigma^2 + \langle K\rangle(\langle K\rangle - 1)$, and therefore

$$q - f^L(0) = \frac{1}{c_L} = \frac{2ye^{-y}}{\sigma_c^2(1-e^{-y})L},$$

introducing the variance at the critical point, σ_c^2.

In order to obtain the probability of surviving, we need to add $1 - q$, which is, using Eq. (3),

$$1 - q = \frac{\langle K\rangle - 1}{f''(1)/2} = 2\frac{\langle K\rangle - 1}{\sigma_c^2}\left[1 - \frac{\langle K\rangle(\langle K\rangle - 1)}{\sigma_c^2}\right] =$$
$$= \frac{2y}{\sigma_c^2 L}\left(1 - \frac{y}{\sigma_c^2 L} - \frac{y^2}{\sigma_c^2 L^2}\right),$$

using also $\langle K\rangle - 1 = y/L$. Therefore, the leading term in $P_{\rm surv}(L)$ turns out to be

$$P_{\rm surv}(L) = 1 - f^L(0) = 1 - q + \frac{1}{c_L}$$
$$= \frac{1}{\sigma_c^2 L}\left(2y + \frac{2e^{-y}y}{1-e^{-y}}\right) = \frac{1}{\sigma_c^2 L}\left(\frac{2ye^y}{e^y-1}\right),$$

which is the same indeed as in the subcritical case.

C. Finite-size scaling law

The previous formula, and its identical replication in the subcritical and critical cases, allows one to write the relationship between the probability of surviving and the control parameter $\langle K\rangle$ in the form of a scaling law,

$$P_{\rm surv}(L) = \frac{1}{L\sigma_c^2}\mathcal{G}[L(\langle K\rangle - 1)], \quad (4)$$

with

$$\mathcal{G}(y) = \frac{2ye^y}{e^y - 1}$$

a universal scaling function, as it is independent of the underlying distribution of K (number of offspring per element). This is valid for large system sizes L and small distances to the critical point, keeping $L(\langle K\rangle - 1) = y$ finite. Notice that

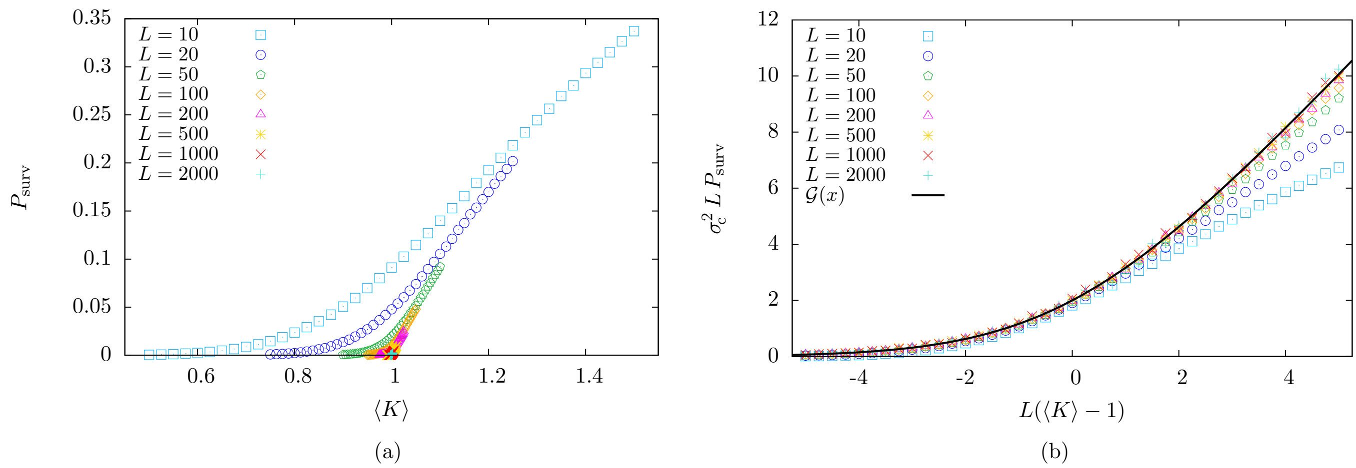

FIG. 1. (Color online) Validity of the finite-size scaling law for the geometric Galton-Watson process. (a) Simulation results for the probability of surviving as a function of the mean number of offspring $\langle K \rangle$ in a Galton-Watson process defined by a geometric distribution of K, given by Prob$[K = k] = p(1-p)^k$, for $k = 0,1,\ldots$ Different values of L show the dependence with system size. Probabilities are estimated from 1000 independent realizations. (b) Same probabilities as a function of $\langle K \rangle - 1$ under rescaling with L and σ_c. The collapse of the curves onto a unique scaling function, given by our theoretical result, is the signature of the validity of the finite-size scaling for large system sizes L and close to the critical point $\langle K \rangle = 1$.

this corresponds to the usual finite-size scaling form, Eq. (1), at zero field, $h = 0$, with exponents

$$\beta = \nu = 1$$

(and metric factor $b = 1$). The only peculiarity is the appearance of an additional metric factor given by σ_c^2. The validity of the finite-size scaling law and its universality is confronted with computer simulations of diverse branching processes in Figs. 1 and 2, with positive results.

From the finite-size scaling law, it is remarkable that keeping fixed the rescaled distance to the critical point, y, the survival probability decays to zero with L as a power

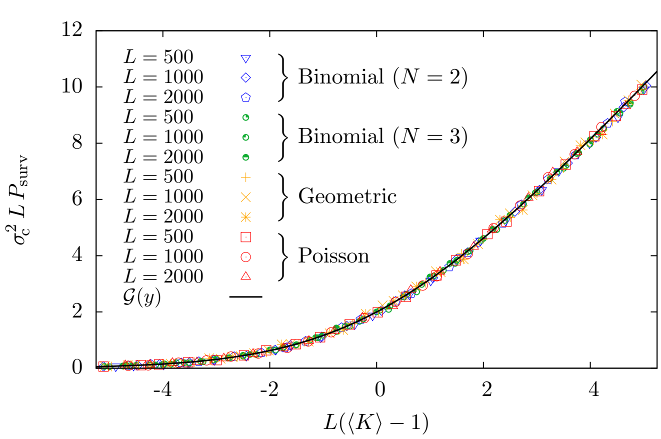

FIG. 2. (Color online) Universality of the finite-size scaling. Extension of Fig. 1(b) to other Galton-Watson processes with different distributions for the number of offspring K. In addition to the geometric case, the binomial distribution, with Prob$[K = k] = N!p^k(1-p)^{N-k}/[k!(N-k)!]$ and $N = 2$ or 3, and the Poisson distribution, with Prob$[K = k] = e^{-\lambda}\lambda^k/k!$, are simulated for different $\langle K \rangle$ and different system sizes L. The collapse of all rescaled curves validates the universality of the scaling exponents and of our scaling function.

law, hyperbolically, see Eq. (4). In contrast, when the absolute distance to the critical point, $\langle K \rangle - 1$, is fixed, the decay is exponential towards the infinite-size probability, $P_{surv}(L) = 1 - q + [f'(q)]^L$, see Refs. [1, p. 16] and [2, p. 38], and Eq. (2).

Interesting information comes from the different limit behaviors of the scaling function,

$$\mathcal{G}(y) \to \begin{cases} -2ye^y & \text{when } y \to -\infty, \\ 2 & \text{when } y \to 0, \\ 2y & \text{when } y \to \infty, \end{cases}$$

which yields

$$P_{surv}(L) \to \begin{cases} 2\sigma_c^{-2}(1 - \langle K \rangle)e^{-L(1-\langle K \rangle)} & \text{for } \langle K \rangle < 1, \\ 2\sigma_c^{-2}L^{-1} & \text{for } \langle K \rangle = 1, \\ 2(\langle K \rangle - 1)/\sigma_c^2 & \text{for } \langle K \rangle > 1. \end{cases}$$

The subcritical and supercritical results have to be understood as holding close to the critical point but for an infinitely large system.

IV. DISCUSSION

As an illustration, let us particularize the finite-size scaling law, Eq. (4), to the case where the number of offspring K of each element is given by a geometric distribution (defined for $K \geqslant 0$), with success parameter p. Then the mean and variance are $\langle K \rangle = (1-p)/p$ and $\sigma^2 = (1-p)/p^2$, the critical point is at $p_c = 1/2$, the variance at p_c is $\sigma_c^2 = 2$, and the finite-size scaling, Eq. (4), transforms into

$$P_{surv}(L) = \frac{1}{2L}\mathcal{G}\left[-4L\left(p - \frac{1}{2}\right)\right]. \tag{5}$$

A more relevant case is when the number of offspring follows a binomial distribution, defined by the outcome of N independent trials with probability of success p in each trial. This leads to a direct connection to the problem of percolation [8–10]. The mean and variance of K are given by $\langle K \rangle = Np$ and $\sigma^2 = Np(1-p)$ (notice that in this case the parameter p has a different interpretation than for the geometrical distribution). The critical point arises then at $p_c = 1/N$, with $\sigma_c^2 = 1 - 1/N$, and the finite-size scaling (4) can be written as

$$P_{\text{surv}}(L) = \frac{1}{L(1-1/N)} \mathcal{G}\left[NL \left(p - \frac{1}{N} \right) \right].$$

Comparing this equation with the corresponding one for the geometric case, Eq. (5), we see that the universality of Eq. (4) becomes "diffused" due to the appearance of two nonuniversal metric factors multiplying L. Clearly, the universal form given by Eq. (4) is preferred. Figure 2 demonstrates this universal finite-size scaling relation for the binomial and the geometric distributions (including as well the Poisson distribution).

It is worthwhile to delve into the analogy between branching processes and percolation. This is established when one considers the latter problem in the Bethe lattice [8,10], and the branching process originates, as we have just mentioned, from a binomial distribution of offspring. The probability p of success in each reproductive trial is just the probability of occupation in site percolation, and the number of trials N is related to the coordination number Z of the Bethe lattice by $Z = N + 1$; then, the survival probability becomes the same as the probability of percolation. Under these conditions, there are still a few differences between branching processes and percolation. In the latter case, in order to establish the probability of percolation, one considers the central site of the Bethe lattice, which can be occupied or not and which is connected (or not) through Z branches [8,10]. In branching processes, this corresponds to the element of the 0 generation, which, in contrast, is always present ($N_0 = 1$) and only gives rise, at most, to $N = Z - 1$ branches (this is the only element that has no ancestor). Therefore, one has to correct for these factors in order to obtain the percolation probability (in the finite-size case) from here, and the way to do it is to take care of the special role of the central site in percolation (in comparison with branching processes).

In order to proceed, we consider the Bethe lattice (with L generations) as composed by the central site plus Z separated branching processes (with $L - 1$ generations, at most) [20]. For each of these branches, the probability of survival, or of percolation, would be $P_{\text{surv}}(L-1)$, if their initial element were occupied for sure, but as it is occupied with probability p the probability of survival becomes $pP_{\text{surv}}(L-1)$. The probability that none of the Z branches survives (i.e., none of them percolates) is $[1 - pP_{\text{surv}}(L-1)]^Z$, and the probability of (at least one) survival starting from the origin is the final probability of percolation [8,10],

$$P_{\text{perco}}(L) = p\{1 - [1 - pP_{\text{surv}}(L-1)]^Z\}$$

(taking into account that the origin is occupied with probability p). Substituting here the universal scaling law, Eq. (4), with $L \simeq L - 1$, and expanding around $p = p_c = 1/N$ and $P_{\text{surv}}(L) = 0$ we get

$$P_{\text{perco}}(L) = Zp_c^2 P_{\text{surv}}(L) = \frac{Z}{(Z-1)(Z-2)L} \mathcal{G}[L(\langle K \rangle - 1)],$$

for percolation in the Bethe lattice, under the condition that the central site is occupied with probability p.

Despite the different definitions of the probability of survival, or of percolation, we see that neither the critical exponents nor the scaling function depend on that. Note that, strictly speaking, $P_{\text{perco}}(L)$ is not the order parameter of the percolation transition but its ensemble-average value, see, for instance, Ref. [21]. Further, it is interesting to point out that the problem of percolation has been extensively studied for random networks [22,23]. The finite-size approach developed in our paper could be extended to such random-network models.

Summarizing, we have found that the second-order phase transition from sure extinction to nonsure extinction in the Galton-Watson branching process fulfills a finite-size scaling law, where the scaling function and the scaling exponents can be exactly derived. If the variance of the distribution of the number of offspring per element is taken into account in the scaling law, this becomes universal, with universal metric factors, in the sense that it is exactly the same for all Galton-Watson branching processes. The results are also valid for percolation in the Bethe lattice, which represents the mean-field limit of the percolation problem.

ACKNOWLEDGMENTS

We are grateful to K. Christensen, with whom we have had discussions about the relation with percolation. R. Garcia-Millan enjoyed a stay at the Centre de Recerca Matemàtica through its Internship Program. The remaining authors acknowledge support from Project No. FIS2012-31324, from Spanish MINECO, and Grant No. 2014SGR-1307 from AGAUR. F. Font-Clos is also grateful to the AGAUR under Grant No. 2012FI_B 00422.

[1] T. E. Harris, *The Theory of Branching Processes* (Springer, Berlin, 1963).
[2] K. B. Athreya and P. E. Ney, *Branching Processes* (Springer-Verlag, Berlin, 1972).
[3] D. G. Kendall, The genealogy of genealogy branching processes before (and after) 1873, Bull. Lond. Math. Soc. **7**, 225 (1975).
[4] M. Kimmel and D. E. Axelrod, *Branching Processes in Biology* (Springer-Verlag, New York, 2002).
[5] I. Pázsit and L. Pál, *Neutron Fluctuations* (Elsevier, Oxford, 2008).
[6] D. Vere-Jones, A branching model for crack propagation, Pure Appl. Geophys. **114**, 711 (1976).
[7] A. Helmstetter and D. Sornette, Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, J. Geophys. Res. B **107**, 2237 (2002).
[8] D. Stauffer and A. Aharony, *Introduction To Percolation Theory*, 2nd ed. (CRC Press, Boca Raton, FL 1994).
[9] A. Bunde and S. Havlin (eds.), *Fractals and Disordered Systems* (Springer, Berlin, 1996).
[10] K. Christensen and N. R. Moloney, *Complexity and Criticality* (Imperial College Press, London, 2005).

[11] S. Zapperi, K. B. Lauritsen, and H. E. Stanley, Self-organized branching processes: Mean-field theory for avalanches, Phys. Rev. Lett. **75**, 4071 (1995).
[12] S. Hergarten, Branching with local probability as a paradigm of self-organized criticality, Phys. Rev. Lett. **109**, 148001 (2012).
[13] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1973).
[14] J. M. Yeomans, *Statistical Mechanics of Phase Transitions* (Oxford University Press, New York, 1992).
[15] H. J. Jensen, *Self-Organized Criticality* (Cambridge University Press, Cambridge, UK, 1998).
[16] G. Pruessner, *Self-Organised Criticality: Theory, Models and Characterisation* (Cambridge University Press, Cambridge, UK, 2012).
[17] A. Corral and F. Font-Clos, in *Self-Organized Criticality Systems* edited by M. J. Aschwanden (Open Academic Press, Berlin, 2013), pp. 183–228.
[18] V. Privman, in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, edited by V. Privman (World Scientific, Singapore, 1990), pp. 1–98.
[19] H. E. Stanley, Scaling, universality, and renormalization: Three pillars of modern critical phenomena, Rev. Mod. Phys. **71**, S358 (1999).
[20] J. W. Essam, Percolation theory, Rep. Prog. Phys. **43**, 833 (1980).
[21] R. Botet and M. Ploszajczak, Exact order-parameter distribution for critical mean-field percolation and critical aggregation, Phys. Rev. Lett. **95**, 185702 (2005).
[22] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
[23] R. Cohen and S. Havlin, *Complex Networks* (Cambridge University Press, Cambridge, 2010).