



Galton–Watson Trees with First Ancestor Interaction

François Dunlop¹ · Arif Mardin²

Received: 9 March 2022 / Accepted: 21 September 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

We consider the set of random Bienaymé–Galton–Watson trees with a bounded number of offspring and bounded number of generations as a statistical mechanics model: a random tree is a rooted subtree of the maximal tree; the spin at a given node of the maximal tree is equal to the number of offspring if the node is present in the random tree and equal to -1 otherwise. We introduce nearest neighbour interactions favouring pairs of neighbours which both have a relatively large offspring. We then prove (1) correlation inequalities and (2) recursion relations for generating functions, mean number of external nodes, interaction energy and the corresponding variances. The resulting quadratic dynamical system, in two dimensions or more depending on the desired number of moments, yields almost exact numerical results. The balance between offspring distribution and coupling constant leads to a phase diagram for the analogue of the extinction probability. On the transition line the mean number of external nodes in generation $n+1$ is found numerically to scale as n^{-2} .

Keywords Random tree · Galton–Watson · Correlation inequalities · FKG · Extinction

1 Introduction

The principal tool of our investigation is a Bienaymé–Galton–Watson (BGW) tree, which is a specific example of those stochastic processes known as “branching processes”. As such, it has found many fruitful applications not only in population dynamics, but also in genetics, nuclear chain reactions, etc. It has an interesting history regarding its origins, which we shall try to outline below. In most textbooks treating the subject, the tree in question is usually called “Galton–Watson tree” because the initial impetus is attributed to the question posed by the British statistician Francis Galton in “Educational Times” in 1873 [1], concerning the

Communicated by Aernout van Enter.

✉ François Dunlop
francois.dunlop@cyu.fr

Arif Mardin
mardin.arif@gmail.com

¹ Laboratoire de Physique Théorique et Modélisation (CNRS, UMR 8089), CY Cergy Paris Université, 95302 Cergy-Pontoise, France

² Nesin Matematik Köyü, Şirince mahallesi, 7, Kayserkaya Sokak, 35920 Selçuk, Izmir, Turkey

possibility of extinction of the names of “noble” families in Britain. After having received several incorrect solutions, he managed to rouse the interest of his mathematician friend Henry William Watson, who posed the problem correctly but his solution was not quite right: He concluded that the probability of extinction is always (i.e., including the supercritical case) equal to 1 [2, 3]. At that time the matter was believed to be settled by the solution of Watson.

Publication of a completely correct solution had to wait until 1930 [4], when the Danish mathematician J.F. Steffensen’s work using contemporary probabilistic tools was published, in a danish mathematical journal. His article being in danish must surely have prevented its immediate recognition. Nevertheless, three years later he published a more comprehensive version of his work in *Annales de l’Institut Henri Poincaré* [5], written in French, thus enabling him to reach a much wider part of the global mathematical community.

Interestingly enough, this is not the whole story as far as historical precedence is concerned. The names of A.K. Erlang, J.B.S. Haldane and, above all, Jules Bienaymé (the same Bienaymé of the famous Bienaymé–Chebychev Inequality of probability theory) should also be mentioned. We shall henceforth focus on Bienaymé’s contribution, and invite the reader to consult David Kendall’s article [6] for a nice exposition of the work of Erlang and Haldane.

Surprisingly, Jules Bienaymé considered the same kind of problem regarding the French aristocracy and famous bourgeois families in France nearly three decades before the appearance of Galton’s famous problem in *Educational Times*. His communication was published in the journal of the *Société Philomatique de Paris* in 1845 [7]. This important discovery was made in 1972 by Heyde and Seneta [8]. What is striking is that although Watson’s solution led to the erroneous conclusion that there would be extinction with probability one even in the supercritical case, Bienaymé had the whole theorem of criticality correctly posed. However, even though a full treatment of the problem for publication has been promised by him in a “*mémoire spécial*”, no trace of it has been found so far. It seems that nobody can be sure if there is yet another and yet even earlier serious attempt to state and prove the same result waiting to be discovered. In our treatment we shall honour the historical precedence and call our tree a Bienaymé–Galton–Watson (BGW) tree. Finally, interested readers are highly recommended to consult the two excellent survey articles by David Kendall [6, 9].

We are interested in large random planar (ordered, labeled) rooted trees, which may have grown by a Markov process such as a Galton Watson process, and are then subject to self-interaction through a Boltzmann weight, with an interaction energy between first ancestor and offspring. The main motivation behind this work is an attempt to understand the following problem. Being given a BGW tree in the usual sense, to what degree can one change the behaviour of the characteristic values such as the mean number of offspring or extinction probabilities by attributing some additional probabilities to those existing between typical parent-offspring pairs: More precisely, what happens if the offspring of a populous family (i.e., parents having at least two children) have populous families themselves? Such relations being nearest-neighbour type due to the tree structure, we introduced an interaction function which increases the likelihood of this sort of outcome. Lattice models of statistical mechanics employ similar ideas, and we tried to adapt its approach to our problem. Somewhat similar questions have been treated in the search of the (discrete) time necessary to find the most recent common ancestor of a large population by Möhle [10].

Two notable features come with our approach. On one side the random tree hierarchical structure yields a fruitful reduction to a dynamical system. On another side the extinction problem of the BGW Markov chain, transferred to the statistical mechanics framework, bears some similarity with the pinning/depinning transition: average tree height bounded or going to infinity in the infinite volume limit. Whence some analogy with the work of Derrida and

Retaux [11] about the depinning transition with disorder on the hierarchical lattice, a toy model of which they solve using the quadratic map of Collet–Eckmann–Glaser–Martin [12].

The outline of the paper is as follows. In Sect. 2 we define the model precisely in terms of a Gibbs measure. In Sect. 3 we design a Markov chain under which the Gibbs measure is invariant. In Sect. 4 we convert random trees to random spin configurations on the maximal tree, where the spin value at any node is -1 if the branch is extinct and the number of offspring otherwise. In Sect. 5 we prove correlation inequalities of the Griffiths and FKG types. In Sect. 6 we establish a recursion relation for generating functions of two parameters, a dynamical system where “time” counts the number of generations and two-dimensional space $\mathbb{R}^2$ represents the two activities associated with single offspring and two or more offspring respectively. In Sect. 7 we extend the dynamical system to $\mathbb{R}^6$ to obtain a recursion for the mean number of external nodes. In Sect. 8 we extend the dynamical system to $\mathbb{R}^4$ to obtain a recursion for the mean energy and to $\mathbb{R}^6$ to obtain a recursion for the variance of the energy. In Sect. 9, in the case where the number of offspring is 0, 1, or 2, we find conditions for the presence of a fixed point in the relevant physical domain. This fixed point is expected to correspond to subcritical or critical trees. In the critical case, running numerically the dynamical system, we find that the mean number of external nodes scales like n^{-2} as $n \nearrow \infty$, contrasting with the constant value 1 in the noninteracting critical BGW case. In Sect. 10, in the case where the number of offspring is 1 or 2, we prove convergence of the specific free energy as $n \nearrow \infty$, and study numerically its dependence upon the coupling constant.

2 Model

For any node i in a tree ω , let $X_i \in \mathbb{N} = \{0, 1, 2, \dots\}$ denote the number of offspring of i . The root is node 0. Nodes are labeled à la Neveu [13]:

$$X_0, X_1, \dots, X_{X_0}, X_{11}, \dots, X_{1X_1}, X_{21}, \dots, X_{2X_2}, \dots \quad (1)$$

The generation of a node is its distance to the root. Except for the root, which belongs to generation 0, the generation of a node equals its number of digits in Neveu notation. It will be denoted $|i|$. The number of generations of a finite tree ω is

$$|\omega| = \max\{|i| : i \in \omega\} \quad (2)$$

We denote $a(i)$ the parent (first ancestor) of i , and $r(i)$ the rank of i within its family, the last digit of its Neveu label.

A probability measure on the set of planar rooted trees ω with no more than n generations is defined as follows. For definiteness we take as “non-interacting” reference measure a BGW probability distribution

$$\mathbb{P}^{GW}(\omega) = \prod_{i \in \omega} p_{X_i} \quad (3)$$

with $p_k = \mathbb{P}^{GW}(X_0 = k)$ of bounded support: $2 \leq K = \max\{k : p_k > 0\} < \infty$. The case of unbounded support will be considered in a forthcoming paper [14]. We denote $\bar{k} = \sum k p_k$. Expectations in $\mathbb{P}^{GW}$ will be denoted $\langle \cdot \rangle_{GW}$ and $\langle A; B \rangle_{GW} = \langle AB \rangle_{GW} - \langle A \rangle_{GW} \langle B \rangle_{GW}$. We recall

$$\left\langle X_0 \right\rangle_{GW} = \bar{k}, \quad \left\langle X_1 + \dots + X_{X_0} \right\rangle_{GW} = \bar{k}^2, \quad \left\langle \sum_{\substack{|i|=m \\ i \in \omega}} X_i \right\rangle_{GW} = \bar{k}^{m+1} \quad (4)$$

Then, given a pair interaction energy

$$\{0, \dots, K\} \times \{0, \dots, K\} \ni (X, Y) \mapsto \varphi(X, Y) \in \mathbb{R} \quad (5)$$

and a boundary condition $X_{a(0)} = x \in \{1, \dots, K\}$ specifying the offspring of a virtual ancestor for the origin, we define a Hamiltonian with first ancestor interaction,

$$H^x(\omega) = \sum_{i \in \omega} \varphi(X_{a(i)}, X_i) \quad (6)$$

and a probability measure on the set of trees ω with at most n generations,

$$\mathbb{P}_n^x(\omega) = \left(\Xi_n^x \right)^{-1} \mathbb{P}^{GW}(\omega) e^{-\beta H^x(\omega)} \quad (7)$$

$$\Xi_n^x = \sum_{\omega} \mathbb{P}^{GW}(\omega) e^{-\beta H^x(\omega)} \quad (8)$$

where $\beta \geq 0$ is the inverse temperature. Expectations in $\mathbb{P}_n^x$ will be denoted $\langle \cdot \rangle_n^x$ and $\langle A; B \rangle_n^x = \langle AB \rangle_n^x - \langle A \rangle_n^x \langle B \rangle_n^x$. We shall be particularly interested in the average total offspring in generation m , or average population in generation $m + 1$, in a tree with n generations,

$$\left\langle \sum_{\substack{|i|=m \\ i \in \omega}} X_i \right\rangle_n^x \quad (9)$$

and in the average total energy

$$\left\langle H^x \right\rangle_n^x \quad (10)$$

The tree represents a hierarchical network, and the nodes are centers of activity. The activity of a node i is measured by its offspring X_i , which may be considered as the number of affiliated centers of activity. The activities of parent and child nodes are expected to be positively correlated, although this could depend upon the type of network. Our basic example will be

$$\varphi(X, Y) = -f(X)g(Y) \quad (11)$$

where f and g are non-negative non-decreasing functions on $\{0, \dots, K\}$. For example

$$\varphi(X, Y) = -1_{X \geq 2} 1_{Y \geq 2} \quad (12)$$

where the indicator function 1_A takes value one if event A is true and zero otherwise.

3 Detailed Balance with Respect to $\mathbb{P}_n^x(\cdot)$

In order to clarify a role that time can play in our study, we now define two discrete-time Markov chains obeying the detailed balance condition with respect to $\mathbb{P}_n^x(\cdot)$. Both Markov chains to be defined below are irreducible and aperiodic in their respective state spaces. The first one is defined as follows. Draw the initial configuration with at most n generations from $\mathbb{P}^{GW}(\cdot)$. Then for $t \geq 0$ take transition probabilities

$$\mathbb{P}^x(\omega^{t+1} = \omega' | \omega^t = \omega) = \mathbb{P}^{GW}(\omega') e^{-\beta(H^x(\omega') - H^x(\omega))_+}, \quad \omega' \neq \omega \quad (13)$$

and $\omega^{t+1} = \omega$ with the complementary probability. The chain may be coupled to an i.i.d. sequence $\bar{\omega}^0, \bar{\omega}^1, \dots, \bar{\omega}^t, \bar{\omega}^{t+1}, \dots$ where each $\bar{\omega}^t$ is drawn independently from $\mathbb{P}^{GW}(\cdot)$.

Take the same initial condition $\bar{\omega}^0 = \omega^0$, and then inductively

$$\omega^{t+1} = \eta_{t+1} \bar{\omega}^{t+1} + (1 - \eta_{t+1}) \omega^t \quad (14)$$

where $\eta_{t+1} = 1$ with probability $e^{-\beta(H^x(\omega^{t+1}) - H^x(\omega^t))_+}$ and $\eta_{t+1} = 0$ with the complementary probability. Detailed balance with respect to (7) is clearly satisfied. Such a dynamics making huge (macroscopic) steps is not very useful.

Our second Markov chain obeying the detailed balance condition with respect to $\mathbb{P}_n^x(\cdot)$ is more like usual Monte Carlo dynamics, and is defined as follows. Draw the initial configuration from $\mathbb{P}^{GW}(\cdot)$. Then at each time step:

- Pick a generation $m \in \{0, 1, \dots, n\}$ randomly according to some probability distribution $\{\lambda_m\}_{m \in \{0, \dots, n\}}$ such that $\lambda_0 > 0$.
- Pick a rank $i_1, \dots, i_m$ à la Neveu, each i_l independently with probability p_{i_l} .
- If the corresponding site $i \notin \omega$, do nothing and exit the time step.
- Flip a fair coin, $\sigma = \pm 1$ with equal probabilities.
- If $X_i + \sigma \notin \{0, \dots, K\}$, do nothing and exit the time step.
- If $\sigma = 1$, with probability p_{X_i+1} let $X'_i = X_i + 1$ and from the additional node, on the right of the X_i already existing nodes, draw a tree with $n - |i|$ generations from $\mathbb{P}_{n-|i|}^{GW}(\cdot)$.
- If $\sigma = -1$, with probability p_{X_i-1} let $X'_i = X_i - 1$ by removing the rightmost node stemming from i and the associated sub-tree.
- If a new X'_i has been defined, accept the new configuration ω' with probability $e^{-\beta(H^x(\omega') - H^x(\omega))_+}$.

Note that $X'_i = X_i + 1 \Rightarrow H^x(\omega') < H^x(\omega)$. Two configurations ω and ω' are connected by one step of the Markov chain if and only if $\exists i \in \omega, i \in \omega'$ such that $X'_i - X_i = \pm 1$ and $\forall j \neq i, j \in \omega \cap \omega', X'_j = X_j$. One can check detailed balance, with $X'_i = X_i + 1$ for definiteness,

$$\frac{\mathbb{P}(\omega \rightarrow \omega')}{\mathbb{P}(\omega' \rightarrow \omega)} = \frac{\mathbb{P}_n^x(\omega')}{\mathbb{P}_n^x(\omega)} = \frac{p_{X'_i}}{p_{X_i}} \frac{\mathbb{P}_{n-|i|}^{GW}(\omega' \setminus \omega)}{\mathbb{P}_{n-|i|}^{GW}(\omega)} e^{-\beta(H^x(\omega') - H^x(\omega))_+} \quad (15)$$

where $\omega' \setminus \omega$ is the configuration ω' restricted to nodes not in ω . The dynamics can be accelerated by dividing all transition rates by $\max\{p_k\}_{0 \leq k \leq K}$.

4 Random Tree as a Spin Model

For any label i , for any planar rooted tree ω , we define $X_i(\omega)$ as the number of offspring of i if $i \in \omega$, and -1 otherwise. Therefore $X_{i_1 \dots i_k} = 0 \Rightarrow X_{i_1 \dots i_k i_{k+1}} = -1 \forall i_{k+1}$. More generally $i_{k+1} > X_{i_1 \dots i_k} \Rightarrow X_{i_1 \dots i_k i_{k+1}} = -1$.

Proposition 1 *Let $n \geq 1$. Let Λ_n be the maximal tree obtained with $X_i = K \forall i$, rooted at 0, with at most n generations. The number of sites $i \in \Lambda_n$ is*

$$|\Lambda_n| = 1 + K + K^2 + \dots + K^n = \frac{K^{n+1} - 1}{K - 1} \quad (16)$$

Let

$$\Omega_n = \{-1, 0, 1, \dots, K\}^{\Lambda_n} \quad (17)$$

By convention, let $p_{-1} = 1$. Fix a boundary condition $X_{a(0)} \in \{1, \dots, K\}$. For $\chi = \{X_i\}_{i \in \Lambda_n} \in \Omega_n$ let

$$\mu^{GW}(\chi) = \prod_{i \in \Lambda_n} p_{X_i} \left(1_{X_{a(i)} \geq r(i)} 1_{X_i \geq 0} + 1_{X_{a(i)} < r(i)} 1_{X_i < 0} \right) \quad (18)$$

where the product counting measure on Ω_n is understood. Recall

$$\mathbb{P}^{GW}(\omega) = \prod_{i \in \omega} p_{X_i} \quad (19)$$

Then there is a bijection $\chi \leftrightarrow \omega$ between the support of μ^{GW} and the set of BGW trees with at most n generations, and $\mu^{GW} \sim \mathbb{P}^{GW}$:

$$\chi \mapsto \omega(\chi), \quad \mathbb{P}^{GW}(\omega(\chi)) = \mu^{GW}(\chi); \quad \omega \mapsto \chi(\omega), \quad \mu^{GW}(\chi(\omega)) = \mathbb{P}^{GW}(\omega) \quad (20)$$

Moreover, let

$$H^{GW}(\chi) = - \sum_{i \in \Lambda_n} \left(1_{X_{a(i)} \geq r(i)} 1_{X_i \geq 0} + 1_{X_{a(i)} < r(i)} 1_{X_i < 0} \right) \quad (21)$$

and for $\lambda > 0$,

$$\mu_\lambda^{GW}(\chi) = Z_\lambda^{-1} \exp(-\lambda H^{GW}) \prod_{i \in \Lambda_n} p_{X_i} \quad (22)$$

where the partition function Z_λ normalizes the probability. Then μ_λ^{GW} converges in distribution to μ^{GW} as $\lambda \rightarrow +\infty$.

In other words, BGW configurations are the ground states of the BGW Hamiltonian (21). Note that by virtue of the boundary condition, we have $X_0 \geq 0$ with probability 1.

Proof

$$\prod_{i \in \Lambda_n} \left(1_{X_{a(i)} \geq r(i)} 1_{X_i \geq 0} + 1_{X_{a(i)} < r(i)} 1_{X_i < 0} \right) \in \{0, 1\} \quad (23)$$

because each factor is 0 or 1. The inverse image of 1 by (23) is the support of μ^{GW} . The set of nodes of $\omega(\chi)$ is the set of sites i such that $X_i \geq 0$. The indicator (23) guarantees that each node has a unique ancestor, which implies that $\omega(\chi)$ is a tree. Conversely the set of sites of $\chi(\omega)$ such that $X_i < 0$ is $\Lambda_n \setminus \omega$. This proves (20). The Hamiltonian (21) takes values $H^{GW}(\chi) \in \{-|\Lambda_n|, -(|\Lambda_n| - 1), \dots, 0\}$. It takes the value $-|\Lambda_n|$ if and only if $\chi = \chi(\omega)$ for some BGW tree ω . There is a gap equal to one relative to the other states. As $\lambda \rightarrow \infty$ the measure concentrates on the ground states, where the Hamiltonian takes the value $-|\Lambda_n|$, which corresponds to (18). For curiosity, an example of χ such that $H^{GW}(\chi) = 0$ is

$$X_i = \begin{cases} -1 & |i| \text{ even} \\ K & |i| \text{ odd} \end{cases} \quad (24)$$

where $|i|$ is the generation of node i . $\square$

Now (5) can be extended to

$$\{-1, 0, \dots, K\} \times \{-1, 0, \dots, K\} \ni (X, Y) \longmapsto \varphi(X, Y) \in \mathbb{R} \quad (25)$$

with $\varphi(X, -1) = \varphi(-1, X) = 0 \forall X$. Given this isomorphism, from now on we'll use freely $\langle \cdot \rangle_{GW}$ and $\langle \cdot \rangle_n^x$ based on either representation of BGW trees, and $\langle \cdot \rangle_{\lambda, GW}$ and $\langle \cdot \rangle_{\lambda, n}^x$ based on (22). The interaction Hamiltonian will be denoted $H^x(\omega)$ or $H^x(\chi)$ according to the context.

For illustration, for any $i \in \Lambda_n$, the probability that the random tree ω includes i , and the mean offspring of i are respectively

$$\mathbb{P}_n^X(\omega \ni i) = \left\langle 1_{X_i \geq 0} \right\rangle_n^X \quad \text{and} \quad \left\langle X_i 1_{X_i \geq 0} \right\rangle_n^X \quad (26)$$

The spin representation can also be viewed as a lattice gas representation with $n_i = X_i + 1$ the number of particles at $i \in \Lambda_n$.

5 Correlation Inequalities

For definiteness we remain with a bounded number of offspring, $X_i \leq K < \infty$, but correlation inequalities can be extended by continuity to any offspring distribution, subject to existence of suitable moments. Also the offspring distribution could depend upon the site $i \in \Lambda_n$, like a random field Ising model.

5.1 Griffiths Inequalities

Following Ginibre [15], let C_n denote the positive cone of multinomials with non-negative coefficients in variables $f(X_i)$ where $|i| \leq n$ and $f(\cdot)$ runs over non-negative non-decreasing functions on $\{-1, 0, \dots, K\}$ with $f(-1) = 0$.

Lemma 1 *Let $n \geq 0$. Let ω and ω' be two independent BGW trees with at most n generations obeying the same probability law $\mathbb{P}^{GW}$. Then for any family $\{f_\alpha, i_\alpha\}_\alpha$ of non-negative non-decreasing functions f_α on $\{-1, 0, \dots, K\}$ with $f_\alpha(-1) = 0$ and node labels i_α with $|i_\alpha| \leq n$, for any choices of $\pm$,*

$$\sum_{\omega, \omega'} \mathbb{P}^{GW}(\omega) \mathbb{P}^{GW}(\omega') \prod_{\alpha} (f_{\alpha}(X_{i_{\alpha}}) \pm f_{\alpha}(X'_{i_{\alpha}})) \geq 0 \quad (27)$$

Moreover for any $F, G \in C_n$,

$$\langle F; G \rangle_{GW} \geq 0 \quad (28)$$

Proof The second assertion is a straightforward consequence of the first, which we prove using (22), where we write

$$\mu_{\lambda}^{GW}(\chi) \mu_{\lambda}^{GW}(\chi') \approx \exp(-\lambda \{H^{GW}(\chi) + H^{GW}(\chi')\}) \prod_i p_{X_i} p_{X'_i} \quad (29)$$

and then

$$\begin{aligned} 1_{X_{a(i)} \geq r(i)} 1_{X_i \geq 0} + 1_{X'_{a(i)} \geq r(i)} 1_{X'_i \geq 0} &= \frac{1}{2} (1_{X_{a(i)} \geq r(i)} + 1_{X'_{a(i)} \geq r(i)}) (1_{X_i \geq 0} + 1_{X'_i \geq 0}) \\ &\quad + \frac{1}{2} (1_{X_{a(i)} \geq r(i)} - 1_{X'_{a(i)} \geq r(i)}) (1_{X_i \geq 0} - 1_{X'_i \geq 0}) \end{aligned} \quad (30)$$

$$\begin{aligned} 1_{X_{a(i)} < r(i)} 1_{X_i < 0} + 1_{X'_{a(i)} < r(i)} 1_{X'_i < 0} &= \frac{1}{2} (1_{X_{a(i)} < r(i)} + 1_{X'_{a(i)} < r(i)}) (1_{X_i < 0} + 1_{X'_i < 0}) \\ &\quad + \frac{1}{2} (1_{X_{a(i)} < r(i)} - 1_{X'_{a(i)} < r(i)}) (1_{X_i < 0} - 1_{X'_i < 0}) \end{aligned} \quad (31)$$

Expanding everything in (27) (29) yields a sum of terms factorized over i , with each factor of the form

$$\sum_{X, X'} p_X p_{X'} \prod_k (1_{X \geq k} + 1_{X' \geq k})^{p_k} \prod_{k'} (1_{X \geq k'} - 1_{X' \geq k'})^{q_{k'}}$$

$$\prod_l (1_{X < l} + 1_{X' < l})^{p'_l} \prod_{l'} (1_{X < l'} - 1_{X' < l'})^{q'_{l'}} \prod_{\alpha_i} (f_{\alpha_i}(X) + f_{\alpha_i}(X')) \prod_{\alpha'_i} (f_{\alpha'_i}(X) - f_{\alpha'_i}(X')) \quad (32)$$

where the sums over X, X' run over $\{-1, 0, \dots, K\}$ and the products over k, k', l, l' run over $\{0, \dots, K\}$, while $p_k, q_{k'}, p'_l, q'_{l'}$ are collections of arbitrary fixed nonnegative integers. The indices α_i, α'_i are for those α which fall on the given site.

The result is zero by symmetry if the number of factors with $-$ sign is odd. Otherwise, up to a factor 2, the summation can be restricted to $X > X'$, where the summand has the sign $(-1)^{\sum_{l'} q'_{l'}}$. Indeed $1_{X < l}$ is a decreasing function while all others are increasing. The product over i then yields a factor $(-1)^{\sum_i \sum_{l'_i} q'_{l'_i}}$, equal to $+1$, because all $q'_{l'}$ factors come in pairs from (31). $\square$

Theorem 1 Let $n \geq 0$ and $x \in \{1, \dots, K\}$. Assume (7) (8) with

$$-H^x(\chi) \in \mathcal{C}_n \quad (33)$$

Then for any $F, G \in \mathcal{C}_n$,

$$\langle F; G \rangle_n^x \geq 0 \quad (34)$$

Moreover $\langle F \rangle_n^x$ is non-decreasing in n and in β and in x and in the coefficient of any term in $-H^x(\cdot)$ as an element of $\mathcal{C}_n$. In particular $\forall 0 \leq m \leq n$, the mean total offspring in generation m obeys

$$\bar{k}^{m+1} \leq \left\langle \sum_{|i|=m} X_i 1_{X_i \geq 0} \right\rangle_n^x \leq K^{m+1} \quad (35)$$

and is increasing in n and has a limit as $n \rightarrow \infty$.

Proof Inequality (34) is a standard consequence of Lemma 1 [15]. The first inequality in (35) is comparison with $\beta = 0$, it follows from (34). The second is trivial. Monotonicity in n is obtained as follows: consider an observable supported in $\{|i| \leq m\}$. Let $n' > n$. Let $\beta = \beta_{|j|}$ depend upon the generation of the link $(j, a(j))$. For $0 \leq s \leq 1$ define

$$\langle \cdot \rangle_{n,n',s} : \quad \beta_{|j|} = s\beta \text{ for } n < |j| < n'$$

Then

$$\langle \cdot \rangle_{n,n',0} = \langle \cdot \rangle_n, \quad \langle \cdot \rangle_{n,n',1} = \langle \cdot \rangle_{n'}$$

Monotonicity in n follows from monotonicity in s , which follows from (34). $\square$

Remark 1 Unlike the Ising model, here there is no spin flip symmetry, hence the restriction to positive f_α, F and G . $\square$

5.2 FKG Inequalities

In the original paper [16], the authors write that one can “extend straightforwardly to more general lattice gases where one allows more than one particle on each site”. Here we give a statement and a proof for our model, following [17].

Theorem 2 (FKG inequality) *Let $n \geq 0$ and $x \in \{1, \dots, K\}$. Assume (5)–(8) (25) with $\forall X, Y, X', Y' \in \{-1, 0, \dots, K\}$*

$$\varphi(X, Y) + \varphi(X', Y') \geq \varphi(X \wedge X', Y \wedge Y') + \varphi(X \vee X', Y \vee Y') \quad (36)$$

Then for any non-decreasing functions $F, G : \Omega_n \rightarrow \mathbb{R}$

$$\langle F; G \rangle_n^x \geq 0 \quad (37)$$

Remark 2 Interaction (11) obeys (36), even without the non-negativity assumption.

Lemma 2 *Let*

$$A_N = \{-1, 0, 1, \dots, K\}^N \quad (38)$$

and let $f_1, f_2, f_3, f_4 : A_N \rightarrow \mathbb{R}_+$ be such that

$$f_1(\chi)f_2(\chi') \leq f_3(\chi \wedge \chi')f_4(\chi \vee \chi') \quad \forall \chi, \chi' \in A_N \quad (39)$$

Then, for any product measure $\mu = \otimes \mu_i$ on A_N ,

$$\langle f_1 \rangle_\mu \langle f_2 \rangle_\mu \leq \langle f_3 \rangle_\mu \langle f_4 \rangle_\mu \quad (40)$$

Proof We first prove the theorem using the lemma with $A_N = \Omega_n$, and then prove the lemma. Let

$$h(\chi) = \prod_{i \in \Lambda_n} \left(1_{X_{a(i)} \geq r(i)} 1_{X_i \geq 0} + 1_{X_{a(i)} < r(i)} 1_{X_i < 0} \right) \quad (41)$$

$$\begin{aligned} f_1(\chi) &= F(\chi)h(\chi)e^{-\beta H^x(\chi)}, & f_2(\chi) &= G(\chi)h(\chi)e^{-\beta H^x(\chi)}, \\ f_3(\chi) &= h(\chi)e^{-\beta H^x(\chi)}, & f_4(\chi) &= F(\chi)G(\chi)h(\chi)e^{-\beta H^x(\chi)} \end{aligned} \quad (42)$$

Applying the lemma gives the theorem. We must check (39), which breaks down into quadruples of factors and for which it suffices that every quadruple obey the inequality. The quadruple containing F and G obeys the inequality by virtue of the monotonicity of F and G . The quadruples from the 2-body interaction, embedded in exponentials, obey the inequality by hypothesis (36). The Galton Watson quadruples obey the inequality because $\forall r, X, Y, X', Y' \in \{-1, 0, 1, \dots, K\}$,

$$\begin{aligned} & \left(1_{X \geq r} 1_{Y \geq 0} + 1_{X < r} 1_{Y < 0} \right) \left(1_{X' \geq r} 1_{Y' \geq 0} + 1_{X' < r} 1_{Y' < 0} \right) \\ & \leq \left(1_{X \wedge X' \geq r} 1_{Y \wedge Y' \geq 0} + 1_{X \wedge X' < r} 1_{Y \wedge Y' < 0} \right) \left(1_{X \vee X' \geq r} 1_{Y \vee Y' \geq 0} + 1_{X \vee X' < r} 1_{Y \vee Y' < 0} \right) \end{aligned} \quad (43)$$

Let us now prove the lemma. Let

$$\chi = (\tilde{\chi}, x), \quad \chi' = (\tilde{\chi}', y), \quad \tilde{\chi}, \tilde{\chi}' \in A_{N-1}, \quad x, y \in \{-1, \dots, K\} \quad (44)$$

$$\tilde{f}_j(\tilde{\chi}) = \langle f_j(\tilde{\chi}, \cdot) \rangle_{\mu_N} = \sum_{x=-1}^K f_j(\tilde{\chi}, x) \mu_N(x), \quad j = 1, 2, 3, 4 \quad (45)$$

We claim that the $\tilde{f}_j$'s obey the hypothesis of the lemma, with $N-1$ in place of N . Iterating N times then proves the lemma. Alternatively one can reason by induction, assuming the lemma up to $N-1$ and applying it to the $\tilde{f}_j$'s. In both cases there remains to prove the claim. We start from the left-hand-side of (39).

$$\begin{aligned} \tilde{f}_1(\tilde{\chi}) \tilde{f}_2(\tilde{\chi}') &= \langle f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', y) \rangle_{\mu_N \otimes \mu_N} \\ &= \langle 1_{x=y} f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', y) \rangle_{\mu_N \otimes \mu_N} \end{aligned}$$

$$+ \langle 1_{x < y} [f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', y) + f_1(\tilde{\chi}, y) f_2(\tilde{\chi}', x)] \rangle_{\mu_N \otimes \mu_N} \quad (46)$$

Given any χ, χ' and $x < y$ let

$$\begin{aligned} a &= f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', y), & b &= f_1(\tilde{\chi}, y) f_2(\tilde{\chi}', x) \\ c &= f_3(\tilde{\chi} \wedge \tilde{\chi}', x) f_4(\tilde{\chi} \vee \tilde{\chi}', y), & d &= f_3(\tilde{\chi} \wedge \tilde{\chi}', y) f_4(\tilde{\chi} \vee \tilde{\chi}', x) \end{aligned} \quad (47)$$

By hypothesis $a \leq c$ and $b \leq c$. Moreover

$$\begin{aligned} ab &= f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', x) f_1(\tilde{\chi}, y) f_2(\tilde{\chi}', y) \\ &\leq f_3(\tilde{\chi} \wedge \tilde{\chi}', x) f_4(\tilde{\chi} \vee \tilde{\chi}', x) f_3(\tilde{\chi} \wedge \tilde{\chi}', y) f_4(\tilde{\chi} \vee \tilde{\chi}', y) = cd \end{aligned} \quad (48)$$

And $a, b \leq c$ with $ab \leq cd$ imply $a + b \leq c + d$. The first term in (46) is bounded as

$$\langle 1_{x=y} f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', y) \rangle_{\mu_N \otimes \mu_N} \leq \langle 1_{x=y} f_3(\tilde{\chi} \wedge \tilde{\chi}', x) f_4(\tilde{\chi} \vee \tilde{\chi}', y) \rangle_{\mu_N \otimes \mu_N} \quad (49)$$

The second term is bounded using $a + b \leq c + d$:

$$\begin{aligned} &\langle 1_{x < y} [f_1(\tilde{\chi}, x) f_2(\tilde{\chi}', y) + f_1(\tilde{\chi}, y) f_2(\tilde{\chi}', x)] \rangle_{\mu_N \otimes \mu_N} \\ &\leq \langle 1_{x < y} [f_3(\tilde{\chi} \wedge \tilde{\chi}', x) f_4(\tilde{\chi} \vee \tilde{\chi}', y) + f_3(\tilde{\chi} \wedge \tilde{\chi}', y) f_4(\tilde{\chi} \vee \tilde{\chi}', x)] \rangle_{\mu_N \otimes \mu_N} \end{aligned} \quad (50)$$

The proof of the lemma is now easily completed. $\square$

6 Recursion for Generating Functions

Let N_n be the number of external nodes of an n -generation tree ω_n :

$$N_n(\omega_n) = \sum_{i \in \omega_n, |i|=n} X_{i_1 \dots i_n} \quad (51)$$

Let $N_n = L_n + Q_n$ where $L_n = L_n(\omega_n)$ denotes the number of external nodes whose parent has one offspring, and $Q_n = Q_n(\omega_n)$ denotes the number of external nodes whose parent has two or more offspring. For $u, v > 0$ let

$$\Xi_n^x(u, v) = \sum_{\omega_n} \mathbb{P}^{GW}(\omega_n) e^{-\beta H^x(\omega_n)} u^{L_n} v^{Q_n} \quad (52)$$

with $H^x(\omega_n)$ as (6) (12), implying

$$\Xi_n^x(u, v) = \Xi_n^2(u, v) \quad \forall x \geq 2$$

The partition function (8) is $\Xi_n^x(1, 1)$. Assume

$$e^{-\beta \varphi(X, Y)} = e^\beta = b > 1 \quad \forall X, Y \geq 2; \quad e^{-\beta \varphi(X, Y)} = 1 \quad \text{whenever } X \text{ or } Y \leq 1 \quad (53)$$

Then

$$\begin{pmatrix} \Xi_0^1(u, v) \\ \Xi_0^2(u, v) \end{pmatrix} = \begin{pmatrix} p_0 + p_1 u + p_2 v^2 + \dots + p_K v^K \\ p_0 + p_1 u + b p_2 v^2 + \dots + b p_K v^K \end{pmatrix} \quad (54)$$

Theorem 3 Let

$$F(u, v) = \begin{pmatrix} p_0 + p_1 u + p_2 v^2 + \dots + p_K v^K \\ p_0 + p_1 u + b p_2 v^2 + \dots + b p_K v^K \end{pmatrix} \quad (55)$$

mapping $[1, \infty) \times [1, \infty)$ into itself and more precisely into $\{1 \leq u < v\}$. Then for $n \geq 0$

$$\Xi_{n+1}^x(u, v) = \Xi_n^x(F(u, v)) = \cdots = \Xi_0^x(F^{(n+1)}(u, v)) \quad (56)$$

(Recursion from the external nodes), and

$$\begin{aligned} \Xi_{n+1}^1(u, v) &= p_0 + p_1 \Xi_n^1(u, v) + p_2 \left(\Xi_n^2(u, v) \right)^2 + \cdots + p_K \left(\Xi_n^2(u, v) \right)^K \\ \Xi_{n+1}^2(u, v) &= p_0 + p_1 \Xi_n^1(u, v) + b p_2 \left(\Xi_n^2(u, v) \right)^2 + \cdots + b p_K \left(\Xi_n^2(u, v) \right)^K \end{aligned} \quad (57)$$

(Recursion from the root)

Denote $(u_n(u, v), v_n(u, v)) = F^{(n)}(u, v)$. Then

$$\Xi_n^1(u, v) = u_{n+1}(u, v), \quad \Xi_n^2(u, v) = v_{n+1}(u, v) \quad (58)$$

When the arguments are not given, we implicitly assume $u = v = 1$, so that $(u_0, v_0) = (1, 1)$ and

$$u_{n+1} = u_{n+1}(1, 1) = \Xi_n^1(1, 1), \quad v_{n+1} = v_{n+1}(1, 1) = \Xi_n^2(1, 1) \quad (59)$$

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} p_0 + p_1 u_n + p_2 v_n^2 + \cdots + p_K v_n^K \\ p_0 + p_1 u_n + b p_2 v_n^2 + \cdots + b p_K v_n^K \end{pmatrix} \quad (60)$$

Proof Let $\omega_n \subset \omega_{n+1}$, and let $i \in N_n$ mean i an external node of ω_n , and similarly $i \in L_n$ and $i \in Q_n$. Then

$$H^x(\omega_{n+1}) = H^x(\omega_n) + \sum_{i \in N_n} \varphi(X_{a(i)}, X_i) \quad (61)$$

$$\begin{aligned} \Xi_{n+1}^x(u, v) &= \sum_{\omega_n} \mathbb{P}^{GW}(\omega_n) e^{-\beta H^x(\omega_n)} \prod_{i \in L_n} (p_0 + p_1 u + p_2 v^2 + \cdots + p_K v^K) \\ &\quad \cdot \prod_{i \in Q_n} (p_0 + p_1 u + b p_2 v^2 + \cdots + b p_K v^K) \\ &= \sum_{\omega_n} \mathbb{P}^{GW}(\omega_n) e^{-\beta H^x(\omega_n)} (p_0 + p_1 u + p_2 v^2 + \cdots + p_K v^K)^{L_n} \\ &\quad \cdot (p_0 + p_1 u + b p_2 v^2 + \cdots + b p_K v^K)^{Q_n} \\ &= \Xi_n^x(p_0 + p_1 u + p_2 v^2 + \cdots + p_K v^K, p_0 + p_1 u + b p_2 v^2 + \cdots + b p_K v^K) \end{aligned}$$

proving (56). Relation (57) is straightforward, and (58) follows from (56) and (54). $\square$

The recursion relation (60) means that the original statistical mechanics problem has been reduced to a discrete time dynamical system in $\mathbb{R}^2$, for which more efficient mathematical and numerical tools are available. While the Monte Carlo simulation is limited to trees with n about a hundred generations, with the usual statistical errors, one can easily run “exactly” the dynamical system up to millions of generations, as shown on Fig. 1. It is also clear from the proof that for $b \neq 1$, there is not a recursion relation involving the diagonal $u = v$ alone.

We have lost the Markov chain of the noninteracting BGW model with n as time. Nevertheless we have gained a dynamical system with n as time.

We shall also use the Fréchet derivative of the map F defined in (55):

$$DF(u, v) = \begin{pmatrix} p_1 & 2p_2 v + \cdots + K p_K v^{K-1} \\ p_1 & 2b p_2 v + \cdots + K b p_K v^{K-1} \end{pmatrix} \quad (63)$$

7 Number of External Nodes

From (52),

$$\begin{aligned} \langle L_n \rangle^x &= \frac{\partial \log \Xi_n^x(u, v)}{\partial u} \Big|_{u=v=1}, & \langle Q_n \rangle^x &= \frac{\partial \log \Xi_n^x(u, v)}{\partial v} \Big|_{u=v=1} \\ \langle L_n \rangle^1 &= \frac{\partial \log u_{n+1}(u, v)}{\partial u} \Big|_{u=v=1} = \frac{1}{u_{n+1}} \frac{\partial u_{n+1}}{\partial u} \Big|_{u=v=1} \end{aligned} \quad (64)$$

$$\langle L_n \rangle^2 = \frac{\partial \log v_{n+1}(u, v)}{\partial u} \Big|_{u=v=1} = \frac{1}{v_{n+1}} \frac{\partial v_{n+1}}{\partial u} \Big|_{u=v=1} \quad (65)$$

$$\langle Q_n \rangle^1 = \frac{\partial \log u_{n+1}(u, v)}{\partial v} \Big|_{u=v=1} = \frac{1}{u_{n+1}} \frac{\partial u_{n+1}}{\partial v} \Big|_{u=v=1} \quad (66)$$

$$\langle Q_n \rangle^2 = \frac{\partial \log v_{n+1}(u, v)}{\partial v} \Big|_{u=v=1} = \frac{1}{v_{n+1}} \frac{\partial v_{n+1}}{\partial v} \Big|_{u=v=1} \quad (67)$$

u_n, v_n and $\partial u_n/\partial u, \partial u_n/\partial v, \partial v_n/\partial u, \partial v_n/\partial v$ at $u = v = 1$ can be computed by induction using

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \\ \partial u_{n+1}/\partial u \\ \partial v_{n+1}/\partial u \\ \partial u_{n+1}/\partial v \\ \partial v_{n+1}/\partial v \end{pmatrix} \Big|_{u=v=1} = \begin{pmatrix} p_0 + p_1 u_n + p_2 v_n^2 + \cdots + p_K v_n^K \\ p_0 + p_1 u_n + b p_2 v_n^2 + \cdots + b p_K v_n^K \\ p_1 \partial u_n/\partial u + 2 p_2 v_n \partial v_n/\partial u + \cdots + K p_K v_n^{K-1} \partial v_n/\partial u \\ p_1 \partial u_n/\partial u + 2 b p_2 v_n \partial v_n/\partial u + \cdots + K b p_K v_n^{K-1} \partial v_n/\partial u \\ p_1 \partial u_n/\partial v + 2 p_2 v_n \partial v_n/\partial v + \cdots + K p_K v_n^{K-1} \partial v_n/\partial v \\ p_1 \partial u_n/\partial v + 2 b p_2 v_n \partial v_n/\partial v + \cdots + K b p_K v_n^{K-1} \partial v_n/\partial v \end{pmatrix} \Big|_{u=v=1} \quad (68)$$

or, using (63),

$$\begin{pmatrix} \partial u_{n+1}/\partial u \\ \partial v_{n+1}/\partial u \end{pmatrix} \Big|_{u=v=1} = DF(u_n, v_n) \begin{pmatrix} \partial u_n/\partial u \\ \partial v_n/\partial u \end{pmatrix} \Big|_{u=v=1} \quad (69)$$

$$\begin{pmatrix} \partial u_{n+1}/\partial v \\ \partial v_{n+1}/\partial v \end{pmatrix} \Big|_{u=v=1} = DF(u_n, v_n) \begin{pmatrix} \partial u_n/\partial v \\ \partial v_n/\partial v \end{pmatrix} \Big|_{u=v=1} \quad (70)$$

It is worth noting that u_n and v_n are mutually coupled but independent of $\partial u_n/\partial u, \partial v_n/\partial u, \partial u_n/\partial v, \partial v_n/\partial v$ at $u = v = 1$, while the latter depend upon u_n, v_n . The recursion starts with

$$\begin{pmatrix} u_0 \\ v_0 \\ \partial u_0/\partial u \\ \partial v_0/\partial u \\ \partial u_0/\partial v \\ \partial v_0/\partial v \end{pmatrix} \Big|_{u=v=1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} u_1 \\ v_1 \\ \partial u_1/\partial u \\ \partial v_1/\partial u \\ \partial u_1/\partial v \\ \partial v_1/\partial v \end{pmatrix} \Big|_{u=v=1} = \begin{pmatrix} 1 \\ p_0 + p_1 + b p_2 + \cdots + b p_K \\ p_1 \\ p_1 \\ 2 p_2 + \cdots + K p_K \\ 2 b p_2 + \cdots + K b p_K \end{pmatrix} \quad (71)$$

Using (64)–(67) yields

$$\langle L_0 \rangle^1 = p_1, \quad \langle L_0 \rangle^2 = \frac{p_1}{p_0 + p_1 + b p_2 + \cdots + b p_K} \quad (72)$$

$$\langle Q_0 \rangle^1 = 2 p_2 + \cdots + K p_K, \quad \langle Q_0 \rangle^2 = \frac{2 b p_2 + \cdots + K b p_K}{p_0 + p_1 + b p_2 + \cdots + b p_K} \quad (73)$$

We have checked $\langle L_n \rangle^x$ and $\langle Q_n \rangle^x$ for $n = 0, 1, 2$ and $x = 1, 2$ and $K = 2, 3$ as above against the Monte Carlo algorithm. The mean total number of external nodes in a tree with

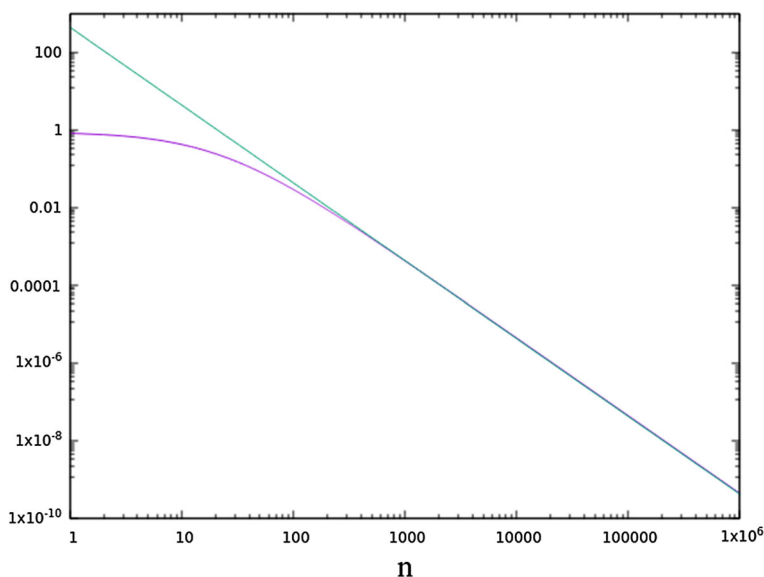


Fig. 1 Mean number of external nodes $\langle N_n \rangle^1$ and n^{-2} fit for $K = 2$, $p_0 = 0.4$, $p_1 = p_2 = 0.3$ and $\beta = \beta_c$ given by equality in (89)

at most n generations, with boundary condition 1 or 2, is given by

$$\langle N_n \rangle^1 = \frac{1}{u_{n+1}} \left(\frac{\partial u_{n+1}}{\partial u} + \frac{\partial u_{n+1}}{\partial v} \right)_{u=v=1}; \quad \langle N_n \rangle^2 = \frac{1}{v_{n+1}} \left(\frac{\partial v_{n+1}}{\partial u} + \frac{\partial v_{n+1}}{\partial v} \right)_{u=v=1} \quad (74)$$

Equation (74) with the recursion (68) was used to generate Fig. 1, further discussed in Sect. 9.

The variances can also be computed by induction, at the expense of four more dimensions to accommodate $\partial^2 u_n / \partial u^2$, $\partial^2 v_n / \partial u^2$, $\partial^2 u_n / \partial v^2$, $\partial^2 v_n / \partial v^2$. We give more detail below for the variance of the energy.

8 Interaction Energy

Let us write (6), (7), (8), (12) as

$$\mathbb{P}_n^x(\omega) = (\Xi_n^x)^{-1} \mathbb{P}^{GW}(\omega) b^{N_{22}} \quad (75)$$

with

$$\Xi_n^x = \sum_{|\omega| \leq n} \mathbb{P}^{GW}(\omega) b^{N_{22}} \quad (76)$$

where $N_{22} = N_{22}(\omega)$ is the number of favoured links: $X_i \geq 2$ and $X_{a(i)} \geq 2$. It depends upon the boundary condition $x = X_{a(0)}$. Then

$$\langle N_{22} \rangle_n^x = (\Xi_n^x)^{-1} b \frac{d}{db} \Xi_n^x \quad (77)$$

or, using (58),

$$\langle N_{22} \rangle_n^1 = \frac{b}{u_{n+1}} \frac{du_{n+1}}{db}, \quad \langle N_{22} \rangle_n^2 = \frac{b}{v_{n+1}} \frac{dv_{n+1}}{db} \quad (78)$$

The relation to energy is given by

$$e^{-\beta H^x(\omega)} = b^{N_{22}} \Rightarrow H^x(\omega) = -\frac{\log b}{\beta} N_{22} \quad (79)$$

The quantities (78) can be computed by recursion:

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \\ du_{n+1}/db \\ dv_{n+1}/db \end{pmatrix} = \begin{pmatrix} p_0 + p_1 u_n + p_2 v_n^2 + \dots + p_K v_n^K \\ p_0 + p_1 u_n + b p_2 v_n^2 + \dots + b p_K v_n^K \\ p_1 \frac{du_n}{db} + 2p_2 v_n \frac{dv_n}{db} + \dots + K p_K v_n^{K-1} \frac{dv_n}{db} \\ p_1 \frac{du_n}{db} + 2b p_2 v_n \frac{dv_n}{db} + \dots + K b p_K v_n^{K-1} \frac{dv_n}{db} \\ \quad + p_2 v_n^2 + \dots + p_K v_n^K \end{pmatrix} \quad (80)$$

starting from (1,1,0,0) at $n = 0$. Or, using (63),

$$\begin{pmatrix} du_{n+1}/db \\ dv_{n+1}/db \end{pmatrix} = DF(u_n, v_n) \begin{pmatrix} du_n/db \\ dv_n/db \end{pmatrix} \quad (81)$$

In particular

$$\begin{pmatrix} u_1 \\ v_1 \\ du_1/db \\ dv_1/db \end{pmatrix} = \begin{pmatrix} 1 \\ p_0 + p_1 + b p_2 + \dots + b p_K \\ 0 \\ p_2 + \dots + p_K \end{pmatrix} \quad (82)$$

$$\langle N_{22} \rangle_0^1 = 0, \quad \langle N_{22} \rangle_0^2 = \frac{b(p_2 + \dots + p_K)}{p_0 + p_1 + b(p_2 + \dots + p_K)} \quad (83)$$

We have checked $\langle N_{22} \rangle_n^x$ for $n = 0, 1, 2$ and $x = 1, 2$ and $K = 2, 3$ as above against the Monte Carlo simulation with the algorithm of Sect. 3.

The algorithm can be extended to the computation of variances:

$$\langle (N_{22})^2 \rangle_n^x - (\langle N_{22} \rangle_n^x)^2 = (\Xi_n^x)^{-1} b^2 \frac{d^2}{db^2} \Xi_n^x + \langle N_{22} \rangle_n^x - (\langle N_{22} \rangle_n^x)^2 \quad (84)$$

$$\langle (N_{22})^2 \rangle_n^1 - (\langle N_{22} \rangle_n^1)^2 = \frac{b^2}{u_{n+1}} \frac{d^2 u_{n+1}}{db^2} + \langle N_{22} \rangle_n^1 - (\langle N_{22} \rangle_n^1)^2 \quad (85)$$

$$\langle (N_{22})^2 \rangle_n^2 - (\langle N_{22} \rangle_n^2)^2 = \frac{b^2}{v_{n+1}} \frac{d^2 v_{n+1}}{db^2} + \langle N_{22} \rangle_n^2 - (\langle N_{22} \rangle_n^2)^2 \quad (86)$$

where the second derivatives are also obtained recursively, adding to (80) the following two lines:

$$\begin{aligned} \frac{d^2 u_{n+1}}{db^2} &= p_1 \frac{d^2 u_n}{db^2} + 2p_2 v_n \frac{d^2 v_n}{db^2} + 2p_2 \left(\frac{dv_n}{db} \right)^2 + \dots \\ &\quad + K p_K v_n^{K-1} \frac{d^2 v_n}{db^2} + K(K-1) p_K v_n^{K-2} \left(\frac{dv_n}{db} \right)^2 \\ \frac{d^2 v_{n+1}}{db^2} &= p_1 \frac{d^2 u_n}{db^2} + b \left[2p_2 v_n \frac{d^2 v_n}{db^2} + 2p_2 \left(\frac{dv_n}{db} \right)^2 + \dots \right. \\ &\quad \left. + K p_K v_n^{K-1} \frac{d^2 v_n}{db^2} + K(K-1) p_K v_n^{K-2} \left(\frac{dv_n}{db} \right)^2 \right] \end{aligned} \quad (87)$$

and starting the map in $\mathbb{R}^6$ from (1, 1, 0, 0, 0, 0).

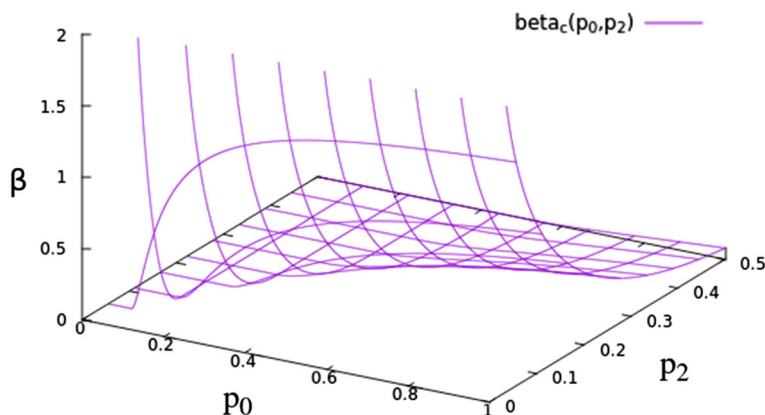


Fig. 2 Critical surface given by equality in (89). The range of (p_0, p_2) to consider is the triangle $p_2 \geq 0$, $p_2 + p_0 \leq 1$, $p_2 - p_0 \leq 0$. The system is expected to be subcritical below the critical surface, going to a fixed point as $n \nearrow \infty$, and supercritical above the critical surface, going to infinity as $n \nearrow \infty$

9 Fixed Point and Phase Diagram

Let $K = 2$. A fixed point (u, v) for (60) reads

$$u = \frac{p_0 + p_2 v^2}{p_0 + p_2}, \quad v^2 \left(p_2(b-1) + \frac{p_2}{p_0 + p_2} \right) - v + \frac{p_0}{p_0 + p_2} = 0 \quad (88)$$

which has a real solution if and only if

$$b-1 \leq \frac{p_0 + p_2}{4p_0 p_2} - \frac{1}{p_0 + p_2} \\ \beta \leq \log \left(1 + \frac{p_0 + p_2}{4p_0 p_2} - \frac{1}{p_0 + p_2} \right) = \beta_c(p_0, p_2) \quad (89)$$

See Fig. 2. For $b \geq 1$ the system (u, v) , started at $(1, 1)$, may only go to a fixed point in $\{v \geq u \geq 1\}$. The first equation in (88) will give a suitable u if $v \geq 1$ has been found. We can therefore restrict our attention to the second equation which may be written as

$$P(v) = v^2 p_2(b-1) + \frac{p_2}{p_0 + p_2} (v-1) \left(v - \frac{p_0}{p_2} \right) = 0 \quad (90)$$

If $p_0 < p_2$, corresponding to a free supercritical BGW, $P(v)$ cannot vanish and there is no fixed point for the interacting system, where $v > 1$. This is consistent with Griffiths inequalities, Theorem 1, implying that $b > 1$ increases the mean number of external nodes, reinforcing supercriticality.

When $\beta = \beta_c(p_0, p_2)$, the fixed point associated with the double root is

$$v = \frac{2p_0}{p_0 + p_2}; \quad u = \frac{p_0}{p_0 + p_2} + \frac{4p_0^2 p_2}{(p_0 + p_2)^3} \quad (91)$$

The phase diagram is linked to the behaviour of the free energy as $n \nearrow \infty$, depending upon b . For boundary condition $x = 1$ or 2, using (16) (58), we find

$$\psi_n^x(b) = -\frac{1}{|\Lambda_n|} \log \Xi_n^x = -\frac{1}{2^{n+1} - 1} \log \Xi_n^x \quad (92)$$

$$\psi_n^1(b) = -\frac{1}{2^{n+1}-1} \log u_{n+1}, \quad \psi_n^2(b) = -\frac{1}{2^{n+1}-1} \log v_{n+1} \quad (93)$$

Clearly, whenever there is convergence to a fixed point, the free energy density (92) (93) vanishes in the limit $n \nearrow \infty$.

When the underlying BGW model is critical or supercritical, $\bar{k} = p_1 + 2p_2 \geq 1$, the interacting model at $\beta > 0$ is supercritical and a corresponding order parameter may be $\rho(b)$ from Theorem 4 below. When the underlying BGW model is subcritical, $p_1 + 2p_2 < 1$, there is $\beta_c(p_0, p_2)$ such that the model is supercritical for $\beta > \beta_c$ and subcritical for $\beta < \beta_c$. Corresponding order parameters may be

$$\lim_{n \nearrow \infty} 2^{-m} \left\langle \sum_{\substack{|i|=m \\ i \in \omega}} X_i \right\rangle_n^x \quad (94)$$

or

$$\lim_{n \nearrow \infty} 2^{-(n-m)} \left\langle \sum_{\substack{|i|=n-m \\ i \in \omega}} X_i \right\rangle_n^x \quad (95)$$

An example for the mean number of external nodes on the critical line is shown on Fig. 1. The behaviour $\langle N_n \rangle^1 \sim n^{-2}$ as $n \rightarrow \infty$ differs from the free critical BGW where $\langle N_n \rangle_{GW} = \bar{k}^n = 1$. A mathematical proof of this behaviour will be given in a forthcoming paper [14].

10 Thermodynamic Limit in the Supercritical Case

We give some results for the simplest interacting supercritical BGW model, namely a model with $K = 2$ and no extinction so that $p_1 + p_2 = 1$, and interaction (12) as before.

Theorem 4 *Let $p_1 + p_2 = 1$ and $0 < p_1 < 1$ and $b > 1$. Recall $(u_n, v_n) = F^{(n)}(1, 1)$ with $F(\cdot)$ as (55). Then*

- (i) $u_n \rightarrow \infty, v_n \rightarrow \infty, v_n/u_n \rightarrow b$ as $n \rightarrow \infty$
- (ii) $(p_2 b v_n)^{2^{-n}}$ strictly increases with n for $n \geq 0$.
- (iii) $\rho(b) = \lim_{n \nearrow \infty} (v_n)^{2^{-n}} = \lim_{n \nearrow \infty} (u_n)^{2^{-n}}$ exists and is non decreasing in b .
- (iv) $\rho(b) = \sup_n (p_2 b v_n)^{2^{-n}} > 1$
- (v) $\sqrt{p_2 b (p_1 + p_2 b)} \leq \rho(b) \leq p_1 + p_2 b$.
- (vi) $\psi(b) = \lim_{n \nearrow \infty} \psi_n^1(b) = \lim_{n \nearrow \infty} \psi_n^2(b) = -\log(\rho(b))$ exists and is non increasing in b .

Remark The Theorem implies, in particular,

$$\begin{aligned} \log \log v_n &= n \log 2 + \log \log \rho + o(n) \\ \log \log u_n &= n \log 2 + \log \log \rho + o(n) \end{aligned} \quad (96)$$

Proof (i) Both u_n and v_n are bigger than they would be with the linear map obtained by replacing v^2 by v and omitting p_0 in (62), with which they would go to infinity exponentially as $n \rightarrow \infty$. Since $v_n \geq u_n$ it follows that $v_n/u_n \rightarrow b$ as $n \rightarrow \infty$.

(ii) We have $v_n > p_2 b v_{n-1}^2$ for $n \geq 1$, or $p_2 b v_n > (p_2 b v_{n-1})^2$, which gives (ii).

(iii–v) Using $\{1 \leq u \leq v\}$ we have

$$v_n \leq (p_1 + p_2 b) v_{n-1}^2 \leq (p_1 + p_2 b) ((p_1 + p_2 b) v_{n-2}^2)^2 \leq \dots$$

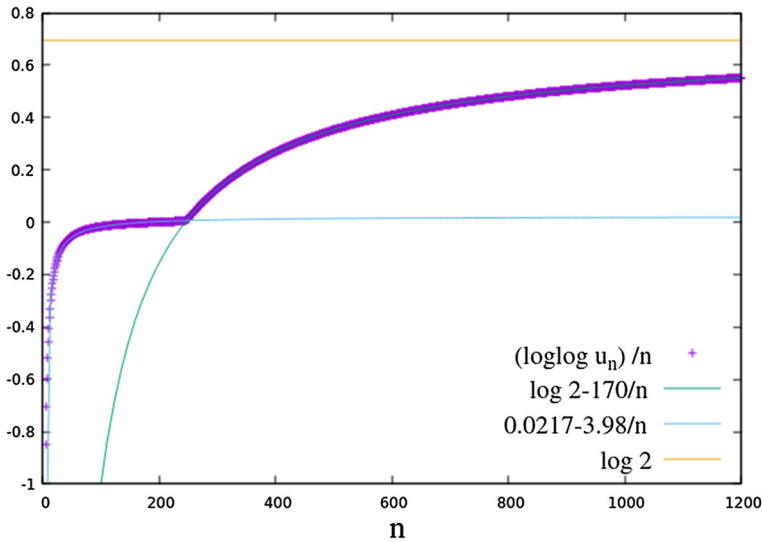


Fig. 3 Numerical values of $\log(\log u_n)/n$ as function of n , for $p_1 = 0.9$, $b = 1.1$, with fits $0.0217 - 3.98/n$ (fitted in range $[10:50]$) and $\log 2 - 170/n$ (fitted in range $[50:1000]$)

$$\begin{aligned}
 &\leq (p_1 + p_2b)^{1+2+4+\dots+2^{m-1}} v_{n-m}^{2^m} \leq \dots \\
 &\leq (p_1 + p_2b)^{1+2+4+\dots+2^{n-2}} v_1^{2^{n-1}} \\
 &= (p_1 + p_2b)^{2^{n-1}-1} (p_1 + p_2b)^{2^{n-1}} = (p_1 + p_2b)^{2^n-1}
 \end{aligned} \tag{97}$$

which can be written as

$$(p_2bv_n)^{2^{-n}} \leq (p_1 + p_2b) \left(\frac{p_2b}{p_1 + p_2b} \right)^{2^{-n}}, \quad n \geq 0 \tag{98}$$

The first claims in (iii–iv) and the upper bound in (v) follow from (i) (ii) and (98). The lower bound in (v) follows from

$$\begin{aligned}
 v_n &> p_2bv_{n-1}^2 > p_2b(p_2bv_{n-2}^2)^2 > \dots > (p_2b)^{1+2+4+\dots+2^{m-1}} v_{n-m}^{2^m} \\
 &= (p_2b)^{2^m-1} (v_{n-m})^{2^m} > \dots > (p_2b)^{2^{n-1}-1} v_1^{2^{n-1}} \\
 &= (p_2b)^{2^{n-1}-1} (p_1 + p_2b)^{2^{n-1}} = (p_2b)^{-1} \left(p_2b(p_1 + p_2b) \right)^{2^{n-1}}
 \end{aligned} \tag{99}$$

(iii) $\rho(b)$ is nondecreasing in b because v_n is a polynomial in b with positive coefficients.

(iv) From the first part of (iv), $\rho(b) > 1$ because $v_n \nearrow \infty$ as $n \nearrow \infty$.

(vi) follows from (iii) and (93). $\square$

If $p_2b(p_1 + p_2b) < 1$ the lower bound in (v) is useless. We examine numerical values, see Fig. 3. For $p_1 = 0.9$, $b = 1.1$ there seems to be a crossover: up to $n \sim 50$ there is a good fit $\frac{1}{n} \log \log u_n \simeq 0.145 - 5.0/n$, while beyond $n \sim 50$ there is a good fit $\frac{1}{n} \log \log u_n \simeq \log 2 - 30.0/n$ (up to $n = 1000$, not shown). This is consistent with (96), with $\log \log \rho \simeq -30.0$. The dependence of ρ or ψ upon b for $p_1 = 0.9$ is sketched in Figs. 4 and 5.

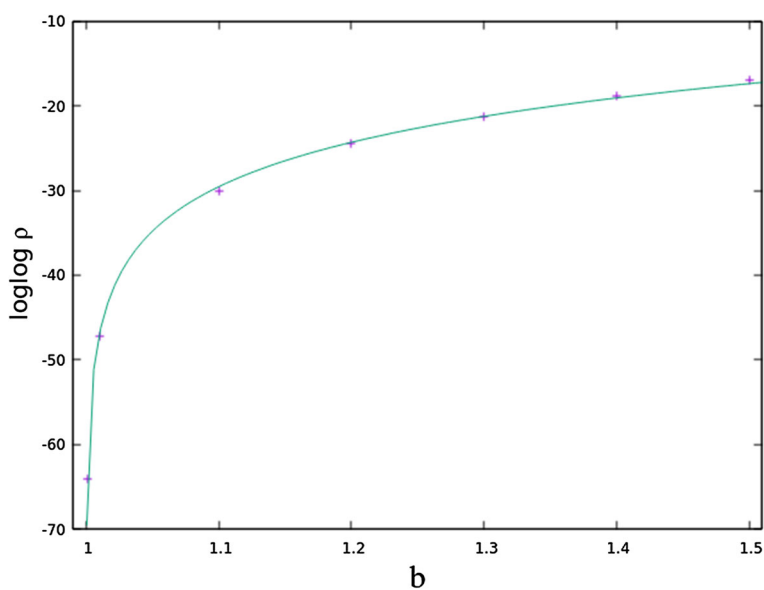


Fig. 4 $\log \log \rho$ as function of b , for $p_1 = 0.9$, with fit $-12.163 + 7.5587 \log(b - 1)$

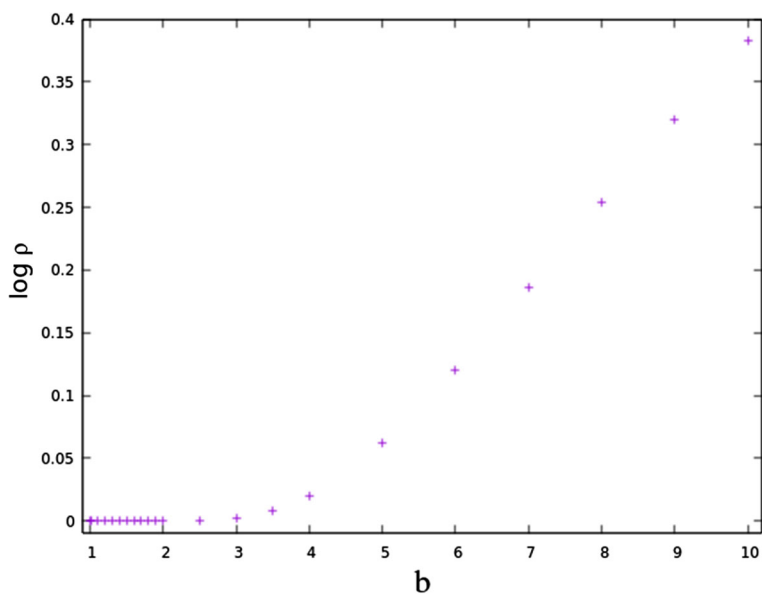


Fig. 5 $\log \rho$ as function of b , for $p_1 = 0.9$

Acknowledgements We would like to express our gratitude to Professor Ali Nesin as well as all the staff of Nesin Mathematics Village for their hospitality and the most pleasant environment for unhindered work they provided during our stay. This work also benefitted at CY Cergy Paris University from the environment of labex MME-DII (Modèles Mathématiques et Economiques de la Dynamique, de l’Incertitude et des Interactions), ANR11-LBX-0023-01.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study. All figures were drawn directly from the given two-dimensional dynamical system and simple use of the gnuplot “fit”.

References

- Galton, F.: Problem 4001. *Educ. Times* **1**, 17 (1873)
- Watson, H.W.: Solution to problem 4001. *Educ. Times* **1**, 115–116 (1873)
- Galton, F., Watson, H.W.: On the probability of extinction of families. *J. R. Anthropol. Inst.* **4**, 138–144 (1874)
- Steffensen, J.F.: On Sandsynligheden for at Afkommet udder, *Matem. Tidsskr. B*, 19–23 (1930)
- Steffensen, J.F.: Deux problèmes du calcul des probabilités. *Ann. Inst. Henri Poincaré* **3**, 319–344 (1933)
- Kendall, D.G.: Branching processes since 1873. *J. Lond. Math. Soc.* **41**, 385–406 (1966)
- Bienaymé, I.J.: De la loi de multiplication et de la durée des familles. *Société Philomatique, Paris, Extraits, Sér.* **5**, 37–39 (1845)
- Heyde, C.C., Seneta, E.: The simple branching process, etc.; a historical note on I. J. Bienaymé *Biometrika* **59**, 680–683 (1972)
- Kendall, D.G.: The genealogy of genealogy: branching processes before (and after) 1873. *Bull. Lond. Math. Soc.* **7**, 225–253 (1975)
- Möhle, M.: The time back to the most recent common ancestor in exchangeable population models. *Adv. Appl. Prob.* **36**, 78–97 (2004)
- Derrida, B., Retaux, M.: The depinning transition in presence of disorder: a toy model. *J. Stat. Phys.* **22**, 268–290 (2014)
- Collet, P., Eckman, J.-P., Glaser, V., Martin, A.: Study of the iterations of a mapping associated to a spin glass model. *Commun. Math. Phys.* **94**, 353–370 (1984)
- Neveu, J.: Arbres et processus de Galton–Watson. *Ann. Inst. Henri Poincaré Probab. Stat.* **22**, 199–207 (1986)
- Collet, P., Dunlop, F., Huillet, T., Mardin, A.: A Gibbsian random tree with nearest neighbour interaction, in preparation
- Ginibre, J.: General formulation of Griffiths’ inequalities. *Commun. Math. Phys.* **16**, 310–328 (1970)
- Fortuin, C.M., Kasteleyn, P.W., Ginibre, J.: Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* **22**, 896103 (1971)
- Friedli, S., Velenik, Y.: *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, Cambridge (2017)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.