The Interaction Light Cone of the Discrete Bak–Sneppen, Contact and other local processes

Tom Bannink, Harry Buhrman, András Gilyén & Mario Szegedy

Volume 176 • Number 3 • Augus

Journal of Statistical Physics

Springer

Springer

Springer

The Interaction Light Cone of the Discrete Bak–Sneppen, Contact and other local processes

Tom Bannink[1] · Harry Buhrman[1] · András Gilyén[1] · Mario Szegedy[2]

Abstract

We consider a class of random processes on graphs that include the discrete Bak–Sneppen process and several versions of the contact process, with a focus on the former. These processes are parametrized by a probability $0 \leq p \leq 1$ that controls a local update rule, and often exhibit a phase transition in this parameter. In general, analyzing properties of the phase transition is challenging, even for one-dimensional chains. In this article we consider a power-series approach based on representing certain quantities, such as the survival probability or the expected number of steps per site to reach the steady state, as a power series in p. We prove that the coefficients of those power series stabilize for various families of graphs, including the family of chain graphs. This phenomenon has been used in the physics community but was not yet proven. We also show that for local events A, B of which the support is a distance d apart we have $\mathrm{cor}(A, B) = O(p^d)$. The stabilization allows for the (exact) computation of coefficients for arbitrary large systems which can then be analyzed using the wide range of existing methods of power series analysis.

Keywords Discrete bak sneppen · Contact process · Power series · Phase transition

✉ Tom Bannink
  tombannink@gmail.com

  Harry Buhrman
  buhrman@cwi.nl

  András Gilyén
  gilyen@cwi.nl

  Mario Szegedy
  mario.szegedy@alibaba-inc.com

1   QuSoft, CWI, University of Amsterdam, Science Park 123, 1098 XG Amsterdam, Netherlands

2   Alibaba Quantum Laboratory, Alibaba Group, Bellevue, WA 98004, USA

🠒 Springer

1 Introduction

In physics, critical behaviour involves systems in which correlations decay as a power law with distance. It is an important topic in many areas of physics and can also be found in stochastic processes on graphs. Often, such systems have a parameter (e.g. temperature) and when it is set to a critical value, the system exhibits critical behaviour. Power series expansion techniques have been used in the physics literature to numerically approximate critical values and associated exponents. It was often observed that the coefficients of such power series stabilize when the system size grows, and we provide a rigorous proof of this for a large class of stochastic processes.

Self-organized criticality is a name common to models where the critical behaviour is present but without the need of tuning a parameter. This concept has been widely studied, see for example [24]. A simple model for evolution and self-organized criticality was proposed by Bak and Sneppen [2] in 1993. In this random process there are n vertices on a cycle each representing a species. Every vertex has a fitness value in [0, 1] and the dynamics is defined as follows. Every time step, the vertex with the lowest fitness value is chosen and that vertex together with its two neighbors get replaced by three independent uniform random samples from [0, 1]. The model exhibits self-organized criticality, as most of the fitness values automatically become distributed uniformly in $[f_c, 1]$ for some critical value $0 < f_c < 1$. This process has received a lot of attention [1,7,20,21], and a discrete version of the process has been introduced in [5]. The model actually appeared earlier in [18] ("model 3") although formulated in a different way and it was also studied in [10] ("CP 3"). In the discrete Bak–Sneppen (DBS) process, the fitness values can only be 0 or 1. At every time step, choose a uniform random vertex with value 0 and replace it and its two neighbors by three independent values, which are 0 with probability p and 1 with probability $1 - p$. The DBS process has a phase transition with associated critical value p_c [4,22].

The Bak–Sneppen process was originally described in the context of evolutionary biology but its study has much broader consequences, e.g., the process was rediscovered in the regime of theoretical computer science [6] as well. To study the limits of a randomized algorithm for solving satisfiability, the discrete Bak–Sneppen process turned out to be a natural process to analyze.

The DBS process is closely related to the so-called contact process (CP), originally introduced in [11]. Sometimes referred to as the basic contact process, this process models the spreading of an epidemic on a graph where each vertex (an individual) can be healthy or infected. Infected individuals can become healthy (probability $1 - p$), or infect a random neighbor (probability p). The contact process has also been studied in the context of interacting particle systems and many variants of it exist, such as a parity-preserving version [14] and a contact process that only infects in one direction [25]. Depending on the particular flavor of the processes, the CP and DBS processes are closely related [4] and in certain cases have the same critical values. The processes are similar in the sense that vertices can be *active* (fitness 0 or infected) or *inactive* (fitness 1 or healthy). The dynamics only update the state in the neighborhood of active vertices with a simple local update rule. In this article we consider a wide class of processes that fit this description, and our proofs are valid in this general setting. We will, however, focus on the DBS process when we present explicit examples.

In this paper we take a power-series approach and represent several probabilities and expectation values as a power series in the parameter p. There is a wealth of physics literature on series analysis in the theory of critical phenomena, see for example [3,12,13] for an overview. Processes typically only have a critical point when the system size is infinite, but numerical simulations often only allow for probing of finite systems. Our main theorem proves, for our general class of processes, that one can extract coefficients of the power series for an arbitrary large system by computing quantities in only a finite system. One can then apply series analysis techniques to these coefficients of the large system. Series expansion techniques have been extensively used for variants of the contact process as well as for closely related directed percolation models [9,14–17,19,25] in order to extract information about critical values and exponents. For example, in [25] the contact process on a line is studied where infection only happens in one direction. In [14] a process is studied where the parity of the number of active vertices is preserved. In both articles, the power series of the survival probability is computed up to 12 terms and used to find estimates for the critical values and exponents. However, in all this work the stabilization of coefficients has been observed[1] but not proven.

Our Main Contribution is a definition of a general class of processes that encapsulates most of the above processes (Definition 1) and an in-depth understanding of the stabilization phenomenon, complete with a rigorous proof (Lemma 5, Theorem 1). The results are illustrated with examples.

Road Map In Sect. 1.1 we will provide two example power series that exhibit the stabilization phenomenon. In Sect. 1.2 we will sketch our results without going into technicalities and explain the intuition behind them, something that we call the Interaction Light Cone. In Sect. 2 we define our general class of processes in more detail and provide our theorems with their proofs. In Sect. 3 we apply our result to the DBS process, and we compute power-series coefficients for several quantities. As an application, we use the method of Padé approximants to extract an estimate for p_c and we estimate a critical exponent that suggests that the DBS process is in the directed percolation universality class.

1.1 Stabilization of Coefficients

There are different ways of defining the DBS process. These definitions are essentially equivalent and only differ slightly in their notion of time, but can be mapped to each other in a straightforward way. For example, one can pick a random vertex in each time step, and only perform an update when the vertex is active, but always count it as a time step. To study infinite-sized systems, one can consider a continuous-time version with exponential clocks at every vertex. Resampling of a vertex and its neighbors happens when the clock of the vertex rings *and* the vertex is active. When calculating time averages, the subtle differences in these definitions can lead to incorrect estimates and should not be overlooked in simulations.

The common in all definitions is that an update is applied if and only if the picked vertex was active. In order to treat the three models equivalently we will count the number of updates instead of time steps. That is, we count the number of times when an active vertex is selected to perform a local update (we count all such occasions even if the update ends up not changing the actual state).

Numerical simulations clearly exhibit the phase transition in the DBS process when p goes from 0 to 1. There is some critical probability p_c such that for $p < p_c$ the active vertices

---

[1] Some work uses stabilization in the number of time steps instead of system size. However, for understanding the critical behavior, system size is the relevant parameter.
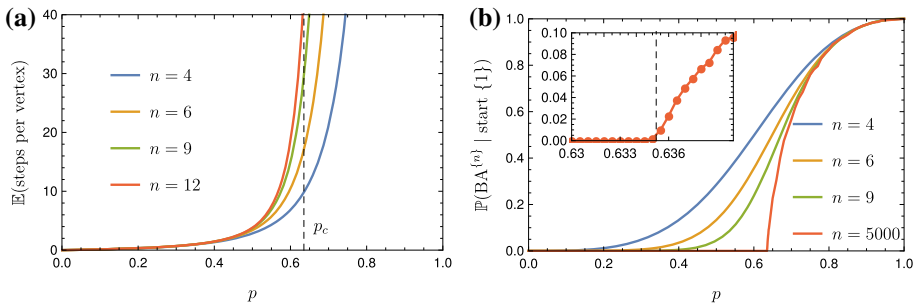
Fig. 1 **a** Plot of $R_{(n)}(p)$, see (1), the expected number of updates per vertex before the all-inactive state is reached, for the DBS process on a cycle with n vertices. The process was started in a random initial state with each vertex being activated independently with probability p. **b** Plot of $S_{[n]}(p)$, see (2), the probability to 'reach' the other side of the system: the DBS process on a non-periodic chain of size n is started with a single active vertex at position 1 (denoted by start {1}) and we plot the probability that vertex n ever becomes active (denoted BA$^{(n)}$) before the all-inactive state is reached. For $n = 5000$ the result was obtained with a Monte Carlo simulation. For the lower n, the results were computed symbolically. The inset shows a zoomed in region of the Monte Carlo data, showing that $p_c \approx 0.635$

quickly die out and the system is pushed toward a state with no active vertices. However for $p > p_c$, the active vertices have the upper hand and dominate the system. This phase transition can clearly be seen in Fig. 1 regarding two different quantities: (a) The expected number of updates per vertex before reaching the all-inactive state on a cycle of length n, after initially activating the vertices i.i.d. randomly with probability p. (b) The probability that the end of a (non-periodic) chain eventually gets activated when the process is started with only one active vertex on the other end.

Let us write these quantities as a power-series in p and in $q = 1 - p$ respectively.

$$R_{(n)}(p) := \frac{1}{n}\mathbb{E}(\text{total updates} \mid \text{start i.i.d.}) = \sum_{k=0}^{\infty} a_k^{(n)} p^k, \tag{1}$$

$$S_{[n]}(q) := \mathbb{P}(\text{vertex } n \text{ becomes active} \mid \text{start } \{1\}) = \sum_{k=0}^{\infty} b_k^{[n]} q^k. \tag{2}$$

We will study these functions in more detail in Sect. 3, where we show, amongst other things, that they are rational functions for each n. For example

$$R_{(4)}(p) = \frac{p(6 - 12p + 10p^2 - 3p^3)}{6(1 - p)^4} = \frac{(1 - q)(1 + q + q^2 + 3q^3)}{6q^4}.$$

Although these quantities only have an operational meaning for $p \in [0, 1]$, we give a plot of such a function over the complex plane, see Fig. 2. The plot shows the poles of $S_{(6)}(p)$, which seem to approach the value p_c on the real line (for larger n see Fig. 6). Similar phenomena can be observed for partition functions in statistical physics. The partition function is usually in the denominator of observable physical quantities, so that its zeros are the poles of such quantities. A classic result on the partition function for certain gasses [26] shows that when an open region around the real axis is free of (complex) zeros, then many physical quantities are analytic in that region and therefore there is no phase transition. Now known as Lee-Yang zeros, they have been widely studied and linked for example to large-deviation statistics [8]. In [23] the hardcore model on graphs with bounded degree is studied, and it is proven that the partition function has zeros in the complex plane arbitrary close to the critical point.
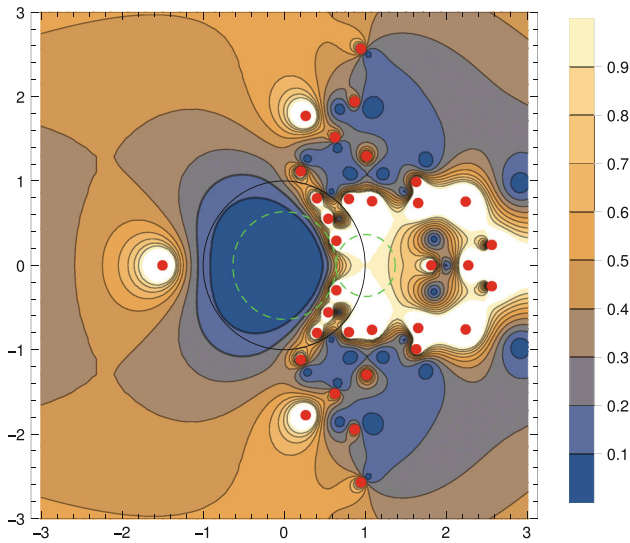
Fig. 2 Plot of the function $|S_{(6)}(p)|$, defined in Eq. (2), over the complex plane with $p = 0$ at the origin. The poles of the function are shown as red dots. The unit circle is shown in black, and the dashed green circles have radius p_c around the origin, and radius $1 - p_c$ around $p = 1$ (Color figure online)

Table 1 Table of the coefficients $a_k^{(n)}$ of the power series defined in Eq. (1)

n	k													
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
3	0	1	**2**	3+1/3	5.0	7.0	9.3	12.0	15.0	18.3	22.0	26.0	30.3	35.0
4	0	1	2	**3+2/3**	6.2	9.7	14.3	20.3	27.8	37.0	48.0	61.0	76.2	93.7
5	0	1	2	3+2/3	**6.4**	10.8	17.3	26.7	39.4	56.5	78.7	106.9	142.2	185.8
6	0	1	2	3+2/3	6.4	**11.0**	18.5	30.0	47.1	71.7	106.0	152.9	215.4	297.4
7	0	1	2	3+2/3	6.4	11.0	**18.7**	31.2	50.8	80.8	125.3	189.7	280.8	407.0
8	0	1	2	3+2/3	6.4	11.0	18.7	**31.4**	52.1	85.0	136.0	213.6	328.9	496.5
9	0	1	2	3+2/3	6.4	11.0	18.7	31.4	**52.3**	86.3	140.7	226.3	358.4	558.4
10	0	1	2	3+2/3	6.4	11.0	18.7	31.4	52.3	**86.5**	142.1	231.6	373.4	594.8
⋮														
14	0	1	2	3+2/3	6.4	11.1	18.8	31.4	52.3	86.5	142.3	233.3	381.2	**621.0**

Although displayed with finite precision, they were computed symbolically

Now we would like to highlight the behaviour of the coefficients $a_k^{(n)}$ and $b_k^{[n]}$. Table 1 and Table 2 show numerical values of the coefficients $a_k^{(n)}$ and $b_k^{[n]}$ respectively.[2]

---

[2] At first sight one is tempted to conjecture that the coefficients $a_k^{(n)}$ are all non-negative and are monotone increasing with n. Unfortunately neither of these conjectures hold since $a_{1114}^{(10)} < 0$. We found this counterexample by observing that the radius of convergence for $R_{(10)}(p)$ is less than 0.96. Since $R_{(10)}(p)$ is bounded on $[0, 0.96]$, this implies that there must be a negative coefficient in its power series.

Table 2 Table of the coefficients $b_k^{[n]}$ of the power series defined in Eq. (2). Although displayed with finite precision, they were computed symbolically

n	k									
	0	1	2	3	4	5	6	7	8	9
3	1	0	**−2**	−2	0	4	6	2	−8	−16
4	1	0	−2	**−4**	−3.5	5.8	22.3	31.3	1.9	−89.1
5	1	0	−2	−4	**−8.3**	−2.5	23.9	76.9	127.9	50.1
6	1	0	−2	−4	−8.3	**−13.7**	2.1	76.6	239.5	422.2
7	1	0	−2	−4	−8.3	−13.7	**−24.6**	19.2	221.4	689.4
8	1	0	−2	−4	−8.3	−13.7	−24.6	**−44.7**	69.28	599.0
9	1	0	−2	−4	−8.3	−13.7	−24.6	−44.7	**−84.2**	197.0
10	1	0	−2	−4	−8.3	−13.7	−24.6	−44.7	−84.2	**−172.3**

A quick look at the table immediately reveals the stabilization of coefficients:

$$a_k^{(n)} = a_k^{(k+1)} \quad \forall n \geq k+1 \quad \text{and} \quad b_k^{[n]} = b_k^{(k+1)} \quad \forall n \geq k+1.$$

Therefore, we now know the first few terms of the power series for arbitrary large systems and we can proceed to use methods of series analysis. By applying the method of Padé approximants, we can numerically estimate $p_c \approx 0.6352$. More details on this can be found in Sect. 3.

1.2 Locality of Update Rule Implies Stabilization

We rigorously prove that the coefficients stabilize, based on an observation that we call the Interaction Light Cone. Let X be a set of vertices, and let L_X be an event that is local on X, meaning that the event depends only on what happens to the vertices in X. For example, when $X = \{v_0\}$ and L_X is the event that vertex v_0 is picked at least r times, then L_X is local on X. In Sect. 2 we will give a more precise definition of local events. We now wish to compare the probability $\mathbb{P}(L_X)$ when the process is initialized in two different starting states, A and A'. When A and A' differ only on vertices that are at least a distance d away from X, then we have

$$\mathbb{P}(L_X \mid \text{start in } A) - \mathbb{P}(L_X \mid \text{start in } A') = O(p^d).$$

By the notation $O(p^d)$ we mean that when this quantity is written as power series in p, then the first $d-1$ terms of the series are zero. It only has non-zero terms of order p^d and higher, i.e., the two probabilities agree on at least the first $d-1$ terms of their power series. This is the essence of the Interaction Light Cone. A vertex that is a distance d away from the set X will only influence probabilities and expectation values of X-local events with terms of order p^d or higher. The intuition behind this is that the probability of a single activation is $O(p)$ and in order for such a vertex to influence the state of a vertex in X, a chain of activations of size d needs to be formed in order to reach X. This observation will also allow us to compare the process on systems of different sizes.

Lemma 1 (Informal version of Lemma 5) *Let G and G' be two graphs and let X be a set of vertices present in both graphs such that the d-neighborhood of X and the local update*

process (where a single update may only affect a vertex and its neighbors) on it is the same in both graphs. Then for any event L_X that is local on X we have

$$\mathbb{P}_G(L_X) = \mathbb{P}_{G'}(L_X) + O(p^d).$$

This idea applies to expectation values as well. Consider the expected number of updates per vertex on a cycle. By translation invariance, we have

$$\frac{1}{n}\mathbb{E}(\text{total updates}) = \mathbb{E}(\#\text{times vertex 1 was updated}),$$

making it a {1}-local quantity. If we add an extra vertex to the cycle, the expectation value only changes by a term of order $O(p^{n/2})$ since the new vertex has distance $n/2$ to vertex 1.

2 Parametrized Local-Update Processes

The class of *parametrized (discrete) local-update processes*, introduced in this section, includes the DBS, the CP and many other natural processes. We prove a general 'stabilization of the coefficients theorem' for them, suggesting the usefulness of the power-series approach for members of the class.

Let $G = (V, E)$ be an undirected graph with vertex set V and edge set E. We consider processes where every vertex of G is either *active* or *inactive*. A *state* is a configuration of active/inactive vertices, denoted by the subset of *active* vertices $A \subseteq V$. For $v \in V$ let us denote by $\Gamma(v)$ the neighbors of v in G including v itself. A local update process in each discrete time step picks a random active vertex $v \in A$ and resamples the state of its neighbors $\Gamma(v)$. If the state is $\emptyset$ (there are no active vertices) then the process stops and all vertices remain inactive afterwards.

Definition 1 (*PLUP - Parametrized local-update process*) We say that M_G is a *parametrized local-update process* on the graph $G = (V, E)$ with parameter $p \in [0, 1]$ if it is a time-independent Markov chain on the state space {inactive, active}^{V} that satisfies the following:

(i) **Initial State** The initial value of a vertex is picked independently from the other vertices. The probability of initializing $v \in V$ as active is a polynomial in p with constant term equal to zero.[3]

(ii) **Selection Dynamics** Each vertex $v \in V$ has a fixed positive weight w_v. A vertex $v \in V$ is selected using one of the three rules[4] below, and if the selected vertex was active, then its neighborhood $\Gamma(v)$ is resampled using the parametrized local-update rule of vertex v (else the state remains unchanged).

    (a) **Discrete-Time Active Sampling** In each discrete time step, an active vertex $v \in A$ is selected with probability $\frac{w_v}{\sum_{u \in A} w_u}$, where A is the current state.

    (b) **Discrete-Time Random Sampling** In each discrete time step, a vertex $v \in V$ is selected with probability $\frac{w_v}{\sum_{u \in V} w_u}$.

    (c) **Continuous-Time Clocks** Every vertex $v \in V$ has an exponential clock with rate w_v. When a clock rings, that vertex is selected, and a new clock is set up for the vertex.

---

[3] The zero constant term is used, for example, in Lemma 4. The independence is used in Lemma 3.

[4] The properties of the selection dynamics are used in the proof of Lemma 3

(iii) **Update Dynamics** The parametrized local-update rule of a vertex $v \in V$ describes a (time-independent) probabilistic transition from state A to A' such that the states only differ on the neighborhood $\Gamma(v)$, i.e., $A \triangle A' \subseteq \Gamma(v)$. The probability P_R of obtaining active vertices $R = A' \cap \Gamma(v)$ is independent of $A \setminus \Gamma(v)$. The probability P_R is a polynomial in p such that for $p = 0$ we get $A' \subsetneq A$ with probability 1, i.e., when any previously inactive vertex becomes active ($|A' \setminus A| > 0$) *or* when $A' = A$ then the constant term in P_R must be zero.[5]

(iv) **Termination** The process terminates when the all-inactive state $\emptyset$ is reached.

With a slight abuse of notation we write $\mathbb{P}_G$ and $\mathbb{E}_G$ for probabilities and expectation values associated to the PLUP M_G, when M_G is clear from context.

Definition 2 (*Local events*) Let $G = (V, E)$ be a (finite) graph and let M_G be a PLUP. Let $S \subseteq V$ be any subset of vertices, and let $v \in V$ be any vertex.

– Let $\mathrm{II}^{(S)}$ be the event that all vertices in S get initialized as inactive.
– Let $\mathrm{RI}^{(S)}$ be the event that all vertices in S remain inactive during the entire process (including initialization).
– Define $\mathrm{BA}^{(S)}$ as the complement of $\mathrm{RI}^{(S)}$: the event that there exists a vertex in S that becomes active at some point during the process, including initialization.
– Let $\#\textsc{Asel}(v)$ be the number of times that v was selected while it was active.
– Let $\#\textsc{Toggles}(v)$ be the number of times that the value of v was changed.

If $S = \{v\}$ we simply use the notation $\mathrm{II}^{(v)}$, $\mathrm{RI}^{(v)}$, and $\mathrm{BA}^{(v)}$ for the above events. We say an event L is **local** on the vertex set S if it is in the sigma algebra generated by the events

$$\mathrm{II}^{(v)}, \ \mathrm{RI}^{(v)}, \ \mathrm{BA}^{(v)}, \ (\#\textsc{Asel}(v) = k), \ (\#\textsc{Toggles}(v) = k) \ : v \in S, 0 \le k < \infty.$$

Lemma 2 (Time equivalence) *The three versions of the selection dynamics of a PLUP, described in property (ii) of Definition 1, are equivalent for local events. That is, for any local event L the probability $\mathbb{P}(L)$ is independent of the chosen selection dynamics in property (ii).*
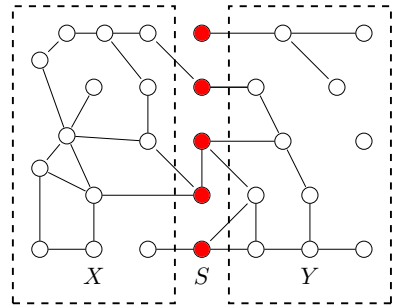
Proof The three selection dynamics only differ in the counting of time, and the presence of self loops in the Markov Chain. The definition of local events only includes events that are independent of the way time is counted. They only depend on which active vertices are selected and the changes to the state of the graph.

It is easy to see that (b) implements the dynamics of (a) via rejection sampling, therefore they give rise to the same probabilities. One can also see that on a finite graph the selection rule (c) induces the same selection rule as (b). This is because the exponential clocks induce a Poisson process at each vertex. The n independent Poisson processes with rates w_v are equivalent to one single Poisson process with rate $W = \sum_{v \in V} w_v$ but where each point of the single process is of type v with probability w_v / W. One can simulate (c) by sampling a time value from an exponential distribution with parameter W and then sampling a random vertex with probability w_v / W (as in (b)). Since the time is not relevant for local events we can ignore the sampled time value and this gives rise to the same probabilities. □

Our lemmas and theorems only concern local events and therefore we can use any one of the three selection dynamics when proving them.

---

[5] The condition $|A' \setminus A| > 0 \implies P_R = O(p)$ is used in the proof of Lemma 4: a fresh activation is at least one power of p so one needs p^k to cover a distance k. The extra condition $A' = A \implies P_R = O(p)$ is used for absolute convergence in Lemma 7 because without it you can have infinitely many paths with a finite power of p. However, we note that this latter condition can be probably slightly relaxed.

Fig. 3 The set S of permanently inactive vertices splits the graph in parts X and Y, rendering them effectively independent. See Lemma 3

Definition 3 (*Induced process*) Suppose that $V' \subseteq V$, then we define the induced process $M_{G'}$ on the induced subgraph $G' = (V', E')$ such that we run the process M_G on G and after each step we deactivate all vertices in $V \setminus V'$. We can then view this as a process on G'. Let L be a local event on V'. We denote the probability of L under the induced process $M_{G'}$ with $\mathbb{P}_{G'}(L)$. Similarly we use the notation $\mathbb{E}_{G'}$ for expectation values induced by the process $M_{G'}$.

It is easy to see that the induced process of a PLUP is also a PLUP.

Definition 4 (*Graph definitions*) Let $G = (V, E)$ be a graph, $S \subseteq V$ be any subset of vertices and $v \in V$ be any vertex.

– Define $G \setminus S$ as the induced subgraph on $V \setminus S$ and $G \cap S$ as the induced subgraph on S.
– Define the d-neighbourhood $\Gamma(S, d)$ of S as the set of vertices that are connected to S with a path of length at most d. In particular $\Gamma(\{v\}; 1) = \Gamma(v)$.
– Define the distant-k boundary $\overline{\partial}(S, k) := \Gamma(S, k) \setminus \Gamma(S, k - 1)$ as the set of vertices lying at exactly distance k from S, and let $\overline{\partial} S := \overline{\partial}(S, 1)$.

The following lemma says that if a set S splits the graph into two disconnected parts, then those two parts become independent under the condition that the vertices in S never become active.

Lemma 3 (Splitting lemma) *Let M_G be a parametrized local-update process on the graph $G = (V, E)$. Let $S, X, Y \subseteq V$ be a partition of the vertices, such that X and Y are disconnected in the graph $G \setminus S$. Furthermore, let L_X and L_Y be local events on X and Y respectively. Then we have (see Fig. 3)*

$$\mathbb{P}_G(\mathrm{RI}^{(S)} \cap L_X \cap L_Y \mid \mathrm{II}^{(S)}) = \mathbb{P}_{G \setminus Y}(\mathrm{RI}^{(S)} \cap L_X \mid \mathrm{II}^{(S)}) \cdot \mathbb{P}_{G \setminus X}(\mathrm{RI}^{(S)} \cap L_Y \mid \mathrm{II}^{(S)}).$$

The condition of initializing S to inactive is present only to prevent counting the initialization probabilities twice. Equivalently we could write the condition only once:

$$\mathbb{P}_G(\mathrm{RI}^{(S)} \cap L_X \cap L_Y) = \mathbb{P}_{G \setminus Y}(\mathrm{RI}^{(S)} \cap L_X) \cdot \mathbb{P}_{G \setminus X}(\mathrm{RI}^{(S)} \cap L_Y \mid \mathrm{II}^{(S)}),$$

and by Bayes rule $\left(\mathbb{P}(L \mid \mathrm{RI}^{(S)}) = \mathbb{P}(L \mid \mathrm{RI}^{(S)} \cap \mathrm{II}^{(S)}) = \frac{\mathbb{P}(L \cap \mathrm{RI}^{(S)} \mid \mathrm{II}^{(S)})}{\mathbb{P}(\mathrm{RI}^{(S)} \mid \mathrm{II}^{(S)})} \right)$ we also have

$$\mathbb{P}_G(L_X \cap L_Y \mid \mathrm{RI}^{(S)}) = \mathbb{P}_{G \setminus Y}(L_X \mid \mathrm{RI}^{(S)}) \cdot \mathbb{P}_{G \setminus X}(L_Y \mid \mathrm{RI}^{(S)}).$$

Proof We will use the 'continuous-time clocks' version of selection dynamics (PLUP property (ii)-c). By Lemma 2 the statement will then hold for all versions. We proceed with a

coupling argument. There are three processes, one on G and the induced ones on $G \setminus Y$ and $G \setminus X$. We couple them by letting all three processes use the same source of randomness. Every vertex in G has an exponential clock that is shared by all three processes, and the randomness used for the local updates for each vertex will also come from the same source. This means that when the clock of a vertex v rings, and the neighborhood $\Gamma(v)$ is equal in different processes, then the update result will also be equal. Now we simply observe that $L_X \cap L_Y \cap \mathrm{RI}^{(S)}$ holds in the G-process if and only if $L_X \cap \mathrm{RI}^{(S)}$ holds in the $(G \setminus Y)$-process and $L_Y \cap \mathrm{RI}^{(S)}$ holds in the $(G \setminus X)$-process. This is because all vertices in S are initialized as inactive (all three probabilities are conditioned on this), so a vertex in S can only be activated by an update from a vertex in X or Y. To check if the event $\mathrm{RI}^{(S)}$ holds, it is sufficient to trace the process up to the first activation of a vertex in S. Before this first activation, anything that happens to the vertices in X only depends on the clocks and updates of vertices in X, and similar for Y. Since S splits X and Y in disconnected parts, these parts can not influence each other unless a vertex in S is activated. Because of the coupling, the evolution of the X vertices in $G \setminus Y$ will be exactly the same as the evolution in G, and similar for Y. Once a vertex in S *does* get activated, the evolution of the three processes is no longer the same but in that case the event $\mathrm{RI}^{(S)}$ does not hold, regardless of any further updates in any system. The clocks and updates of each vertex are independent sources of randomness, and when $\mathrm{RI}^{(S)}$ holds then all the randomness of the S vertices is ignored. Therefore the probability of $\mathrm{RI}^{(S)}$ in the $(G \setminus Y)$-process and $(G \setminus X)$-process depends only on independent random variables corresponding to the vertices in X and Y respectively, and we get the required equality. □

2.1 Interaction Light Cone Results

Now we present the results that exhibit the interaction light cone. The intuition is that if two vertices have distance d in the graph, then the only way they can affect each other is that an interaction chain is forming between them, meaning that every vertex gets activated at least once in between them.

When we write $f(p) = O(p^k)$ for some function f then we mean the following: $f(p)$ is analytic in a neighborhood of 0 and when $f(p)$ is written as a power-series in p, i.e., $f(p) = \sum_{i=0}^{\infty} \alpha_i p^i$, then $\alpha_i = 0$ for $0 \leq i \leq k-1$.

Lemma 4 *Let M_G be a parametrized local-update process on the graph G with vertex set V. Let $\{X_1, \ldots, X_k\}$ be a collection of disjoint vertex subsets $X_i \subseteq V$ and let E be an event. If $E \subseteq \bigcap_i \mathrm{BA}^{(X_i)}$, then $\mathbb{P}(E) = O(p^k)$. Furthermore if $S \subseteq V$ then also $\mathbb{P}(E \mid \mathrm{II}^{(S)}) = O(p^k)$.*

When the event E holds, each set X_i contains a vertex that becomes active, and by PLUP properties (i) and (iii) any activation (either during initialization or later) is $O(p)$. Therefore the probability of E is of order p^k or higher. We give the full proof in Appendix B.

Lemma 5 (Graph surgery) *Let M_G be a parametrized local-update process on the graph $G = (V, E)$. If $X, Y \subseteq V$, $X \cap Y = \emptyset$ and L_X is a local event on X, then*

$$\mathbb{P}_G(L_X) - \mathbb{P}_{G \setminus Y}(L_X) = O(p^{d(X,Y)}).$$

Proof We can assume without loss of generality, that $X \neq \emptyset \neq Y$, otherwise the statement is trivial. Also we can assume without loss of generality that $d(X, Y) \leq \infty$, i.e., X, Y are in the same connected component of G, otherwise we can use Lemma 3 with $S = \emptyset$.

The proof goes by induction on $d(X, Y)$. For the base case, $d(X, Y) = 1$, first note that when $p = 0$, the process initializes everything to inactive by property (i). Depending
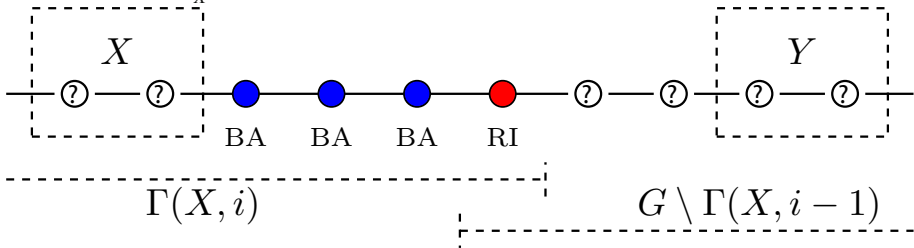
on whether this atomic event is included in L_X, the probability $\mathbb{P}(L_X)$ for $p = 0$ (i.e. the constant term) is either 0 or 1 and independent of the graph.

Now we show the inductive step, assuming we know the statement for d, and that $d(X, Y) = d + 1$. First we assume, that $\mathrm{RI}^{(X)} \subseteq \overline{L_X}$, i.e., $L_X \subseteq \mathrm{BA}^{(X)}$. Define

$$L_X^i := L_X \cap \mathrm{RI}^{(\overline{\partial}(X,i))} \cap \bigcap_{j \in [i-1]} \mathrm{BA}^{(\overline{\partial}(X,j))} \quad \text{for} \ \ i \in [d],$$

$$L_X^{d+1} := L_X \cap \bigcap_{j \in [d]} \mathrm{BA}^{(\overline{\partial}(X,j))}.$$

When L_X^i holds, all vertices at distance i remain inactive, but for all $j \le i - 1$ there exists a vertex at distance j that become active. These events form a partition $L_X = \dot{\bigcup}_{i \in [d+1]} L_X^i$. Below we depict L_X^i graphically:

It is easy to see that for all $i \in [d + 1]$ we have $L_X^i \subseteq \mathrm{BA}^{(X)} \cap \bigcap_{j \in [i-1]} \mathrm{BA}^{(\overline{\partial}(X,j))}$, and therefore by Lemma 4 we get

$$\mathbb{P}_G(L_X^i) = O(p^i), \quad \text{and} \quad \mathbb{P}_G(L_X^i \mid \mathrm{II}^{(\overline{\partial}(X,i))}) = O(p^i). \tag{3}$$

Now we use, for all $i \in [d]$, the Splitting lemma 3 with $S = \overline{\partial}(X, i)$ to split $\Gamma(X, i - 1)$ from $G \setminus \Gamma(X, i)$. We get

$$
\begin{aligned}
\mathbb{P}_G(L_X^i) &= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \mathrm{II}^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{G \setminus \Gamma(X,i-1)}(\mathrm{RI}^{(\overline{\partial}(X,i))}) && \text{(by Lemma 3)} \\
&= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \mathrm{II}^{(\overline{\partial}(X,i))}) \cdot \left(\mathbb{P}_{G \setminus Y \setminus \Gamma(X,i-1)}(\mathrm{RI}^{(\overline{\partial}(X,i))}) + O(p^{d+1-i}) \right) && \text{(by induction)} \\
&= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \mathrm{II}^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{G \setminus Y \setminus \Gamma(X,i-1)}(\mathrm{RI}^{(\overline{\partial}(X,i))}) + O(p^{d+1}) && \text{(by Equation (3))} \\
&= \mathbb{P}_{G \setminus Y}(L_X^i) + O(p^{d+1}) && \text{(by Lemma 3)} \\
&= \mathbb{P}_{G \setminus Y}(L_X^i) + O(p^{d(X,Y)}). && (4)
\end{aligned}
$$

Therefore

$$
\mathbb{P}_G(L_X) \overset{(3)}{=} \sum_{i \in [d]} \mathbb{P}_G(L_X^i) + O(p^{d(X,Y)}) \overset{(4)}{=} \sum_{i \in [d]} \mathbb{P}_{G \setminus Y}(L_X^i) + O(p^{d(X,Y)})
$$

$$
\overset{(3)}{=} \mathbb{P}_{G \setminus Y}(L_X) + O(p^{d(X,Y)}).
$$

We finish the proof by observing that $\mathrm{RI}^{(X)}$ is an atomic event of the sigma algebra of the local events of X, so if $\mathrm{RI}^{(X)} \not\subseteq \overline{L_X}$, then we necessarily have $\mathrm{RI}^{(X)} \subseteq L_X$. Therefore we can use the above proof with $C_X := \overline{L_X}$ and use that $\mathbb{P}(L_X) = 1 - \mathbb{P}(C_X)$. $\qquad \square$

Corollary 1 (Decay of correlations) *Let M_G be a parametrized local-update process on the graph $G = (V, E)$. If $X, Y \subseteq V$ and L_X, L_Y are local events on X and Y respectively, then*

$$\text{Cov}(L_X, L_Y) = \mathbb{P}_G(L_X \cap L_Y) - \mathbb{P}_G(L_X)\mathbb{P}_G(L_Y) = O(p^{d(X,Y)-1}), \tag{5}$$

and

$$\mathbb{P}_G(\text{BA}^{(X)} \cap \text{BA}^{(Y)}) - \mathbb{P}_G(\text{BA}^{(X)})\mathbb{P}_G(\text{BA}^{(Y)}) = O(p^{d(X,Y)+1}). \tag{6}$$

The proof of this lemma is analogous to the proof of Lemma 5 and can be found in Appendix A.

In order to state our general result about the stabilization of the coefficients in the power series we define a notion of isomorphism between different PLUPs.

Definition 5 (*PLUP isomorphism*) We say that the PLUPs M_G and $M_{G'}$ are isomorphic with the fixed sets X, X' if there is a graph isomorphism $i : G \to G'$ such that $i(X) = X'$. Moreover, the probability of transitioning in one step from a state A to A' is preserved under the isomorphism:

$$\mathbb{P}_G(A \text{ is transformed to } A') = \mathbb{P}_{G'}(i(A) \text{ is transformed to } i(A')),$$

and similarly the probability of initializing to a particular state A is preserved:

$$\mathbb{P}_G(\text{graph state is initially } A) = \mathbb{P}_{G'}(\text{graph state is initially } i(A')).$$

We denote such an isomorphism relation by

$$M_G \overset{X}{\underset{X'}{\sim}} M_{G'}.$$

Now we define convergent families of PLUPs. Our requirements for such a family of processes imply that the underlying graphs converge locally, in the neighborhood of a fixed point, to a common graph limit, also called graphing, therefore justifying the term "convergent". Examples of convergent families of PLUPs include DBS and CP on tori of any dimension, when the limit graphing is just the infinite grid. Less regular examples are also included, such as toroid ladder graphs or discrete Möbius stripes of fixed width.

Definition 6 (*Convergent family of PLUPs*) We say a family $\{(M_{G_j}, v_j) : j \in \mathbb{N}\}$ of rooted PLUPs is convergent, if for all $d \in \mathbb{N}$ and for all $j, k \geq d$ we have $M_{\Gamma_{G_j}(\{v_j\},d)} \overset{v_j}{\underset{v_k}{\sim}} M_{\Gamma_{G_k}(\{v_k\},d)}$.

We are ready to state our generic result about the stabilization of coefficients.

Theorem 1 (Power series stabilization) *Suppose that $\{(M_{G_j}, v_j) : j \in \mathbb{N}\}$ is a convergent family of rooted PLUPs, then the coefficients of the power series of $R_{G_i} = \mathbb{E}_{G_i}(\#\text{ASEL}(v_i))$ stabilize. In particular, $R_{G_i}(p) = R_{G_j}(p) + O(p^{\min(i,j)+1})$*

Note that for vertex-transitive graphs, this implies $R_{G_i} = \frac{1}{|G_i|}\mathbb{E}_{G_i}(\text{total updates})$ stabilizes.

Proof Let $d = \min(i, j)$, then

$$
\mathbb{E}_{G_i}(\#\text{ASEL}(v_i)) = \sum_{k \geq 0} k \cdot \mathbb{P}_{G_i}(\#\text{ASEL}(v_i) = k)
$$

$$
= \sum_{k \geq 0} k \cdot \mathbb{P}_{G_i \cap \Gamma_{G_i}(v_i, d)}(\#\text{ASEL}(v_i) = k) + O(p^{d+1}) \quad \text{(by Lemma 5)}
$$

$$
= \sum_{k \geq 0} k \cdot \mathbb{P}_{G_j \cap \Gamma_{G_j}(v_j, d)}(\#\text{ASEL}(v_j) = k) + O(p^{d+1})
$$

$$
\left(M_{G_i \cap \Gamma_{G_i}(v_i, d)} \overset{v_i}{\underset{v_j}{\simeq}} M_{G_j \cap \Gamma_{G_j}(v_j, d)} \right)
$$

$$
= \sum_{k \geq 0} k \cdot \mathbb{P}_{G_j}(\#\text{ASEL}(v_j) = k) + O(p^{d+1}) \quad \text{(by Lemma 5).}
$$

In Lemma 8 in Appendix B, we prove that these types of sums are absolutely convergent for small enough p. Therefore the equality holds when the left- and right-hand side are considered as a power series in p. □

3 The Discrete Bak–Sneppen Process

In Sect. 1.1 we introduced two quantities that exhibit a phase transition in the DBS process. We saw that the coefficients of their power series stabilize. In this section we will look at them in more detail.

3.1 Notation

We denote by M_G the DBS process on the graph $G = (V, E)$. With a slight abuse of notation we also denote by M_G the leaking transition matrix of this time-independent Markov Chain, where the row and column that correspond to the all-inactive configuration are set to zero. We will index vectors (and matrices) by sets $A \subseteq V$, where A is the set of active vertices, as in Sect. 2. We will denote probability row vectors by $\rho \in \mathbb{R}^{2^n}$ so that $\rho \cdot M_G$ is the state of the system after one time step (one update). Setting the all-inactive row and column to zero corresponds to the property that for every $A \subseteq V$ we have $(M_G)_{\emptyset, A} = (M_G)_{A, \emptyset} = 0$. We will use the notation $M_{(n)}$ for the matrix of the process on the cycle of length n and $M_{[n]}$ for the process on the chain (not periodic) of length n. In both case we identify vertices with $V := [n] = \{1, 2, \ldots, n\}$.

3.2 Expected Number of Resamples Per Site

The first quantity of interest is the expected number of updates per vertex to reach the all-inactive state. Consider the DBS process on the cycle of length n. We start the process by letting each vertex be active with probability p and inactive with probability $1 - p$, independently for each vertex. Denote this initial state by $\rho^{(0)}$, so its components have values $\rho_A^{(0)} = p^{|A|}(1 - p)^{n - |A|}$. Let J be the vector with all entries equal to 1, except for the entry of the all-inactive state which is zero. Then $\rho^{(0)} \cdot M_{(n)}^k \cdot J^T$ is the probability that after exactly k updates there is at least one active vertex, i.e. the *all-inactive* state is reached after at least $k + 1$ updates, starting from $\rho^{(0)}$. Now define $R_{(n)}(p)$ as the expected number of updates per
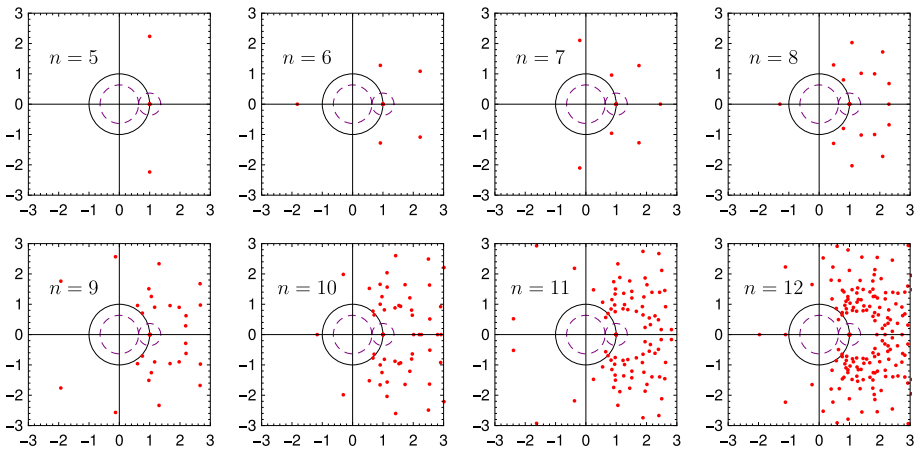
Fig. 4 Location of the poles of $R_{(n)}(p)$ on the complex plane for different n. The black circle is the complex unit circle and the dashed circles have radius p_c around $p = 0$ and $1 - p_c$ around $p = 1$. There is always a pole at $p = 1$ because $R_{(n)}(1)$ is always infinite

vertex, before reaching the all-inactive state:

$$R_{(n)}(p) = \frac{1}{n} \sum_{k=1}^{\infty} k \cdot \mathbb{P}(\text{reach all-inactive in exactly } k \text{ updates})$$

$$= \frac{1}{n} \sum_{k=1}^{\infty} \mathbb{P}(\text{reach all-inactive in } k \text{ updates or more})$$

$$= \frac{1}{n} \sum_{k=1}^{\infty} \rho^{(0)} \cdot M_{(n)}^{k-1} \cdot J^T \tag{7}$$

$$= \frac{1}{n} \rho^{(0)} \cdot (\text{Id} - M_{(n)})^{-1} \cdot J^T \quad \text{(by the geometric series)}$$

$$= \frac{P_{(n)}(p)}{P'_{(n)}(p)}, \tag{8}$$

where $P_{(n)}$, $P'_{(n)}$ are polynomials as can be seen by using Cramer's rule for matrix inversion. Therefore we can conclude that $R_{(n)}(p)$ is a rational function. For small n we can compute $R_{(n)}(p)$ by symbolically inverting the matrix $\text{Id} - M_{(n)}$, which is how we obtained the coefficients in Table 1. For $n \geq 9$ we computed the matrix inverse for rational values of p exactly, and then computed the rational function using Thiele's interpolation formula.

3.2.1 The Power-Series of $R_{(n)}(p)$

As we have seen in the previous subsection, $R_{(n)}(p)$ is a rational function. Since a rational function is analytic, and $R_{(n)}(p)$ has no pole at $p = 0$ (it actually takes value 0), we can write it as

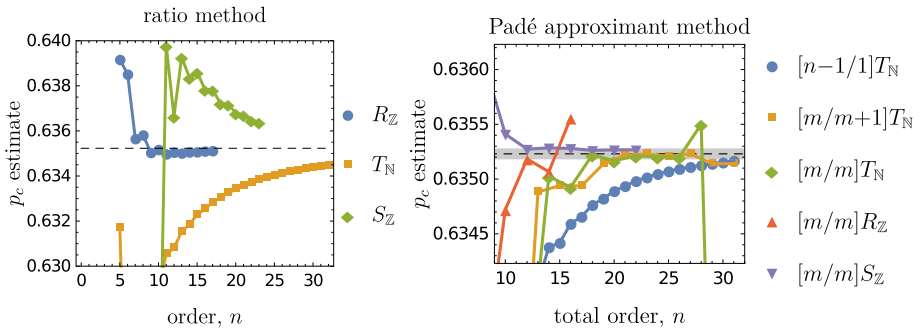$$R_{(n)}(p) = \sum_{k=0}^{\infty} a_k^{(n)} p^k, \tag{9}$$

Fig. 5 Estimates for p_c based on the two methods. On the horizontal axis, n is the number of power-series coefficients used for the estimate. The function $R_{\mathbb{Z}}$, $T_{\mathbb{N}}$ and $S_{\mathbb{Z}}$ are defined in the text below Conjecture 1. The numbers $[m, m']$ (with $m + m' = n$) refer to the degree of the numerator and denominator respectively of the rational functions used in the Padé approximant method. The gray shaded region shows our estimate $p_c = 0.63523 \pm 0.00005$

where the (non-zero) radius of convergence of the above power series equals the absolute value of the closest pole of $R_{(n)}(p)$ to 0. In order to get some intuition about the radius of convergence we plotted the location of the poles of $R_{(n)}(p)$ on the complex plane in Fig. 4. For $n = 10$ there is a pole at a point with absolute value ≈ 0.9598, hence $R_{(10)}(p)$ has a radius of convergence strictly smaller than 1 even though the rational function $R_{(n)}(p)$ is well-defined for all $p \in [0, 1)$.

As was shown in Sect. 1.1, Table 1, the coefficients $a_k^{(n)}$ stabilize as n grows. This is proven by Theorem 1, since the family of DBS processes on the cycles, indexed by n, is a convergent family of PLUPs. The theorem only guarantees the stabilization for $n > 2k$ since going from a cycle of size n to $n + 1$ adds a vertex at a distance $n/2$ to any fixed vertex. In the table, however, we saw that the stabilization already holds for $n \geq k + 1$. In Appendix D we prove this more precise version of the stabilization that holds for cycles. We define the 'stabilized' coefficients $a_k^{(\infty)} := a_k^{(k+1)}$. We then define $R_{\mathbb{Z}}(p) = R_{(\infty)}(p) = \sum_{k=0}^{\infty} a_k^{(\infty)} p^k$ and make the following conjecture.

Conjecture 1 (Radius of convergence) *The radius of convergence of $R_{(\infty)}(p)$ is equal to the critical probability p_c of the DBS process.*

In Appendix B we explain an alternative method to compute coefficients of the $R_{(\infty)}(p)$ power series (see the text below Lemma 9). As an application, we can apply known methods of series analysis. For example, Fig. 5 shows estimates for p_c using the ratio method and the Padé approximant method. For details on these methods, see for example [12]. The ratio method can be used to estimate the critical value when the singularity that determines the radius of convergence is at p_c, i.e. there are no other singularities closer to the origin, which is what we suggest in Conjecture 1. The figure also shows estimates based on the power-series coefficients of the functions $T_{\mathbb{N}}$ and $S_{\mathbb{Z}}$. The function $T_{\mathbb{N}}$ is the expected number of total updates on a semi-infinite chain with one end, with a single active vertex at that end as a starting state. This series is included because we can compute more terms for it. The function $S_{\mathbb{Z}}$ is the probability of survival on the infinite line with a single active vertex as a starting state. This is a series in $q = 1 - p$ and it is included because other work studies the equivalent function for the contact process and this allows for comparison of critical exponents [9]. The Padé approximant method suggests that the critical value is $p_c \approx 0.63523 \pm 0.00005$,
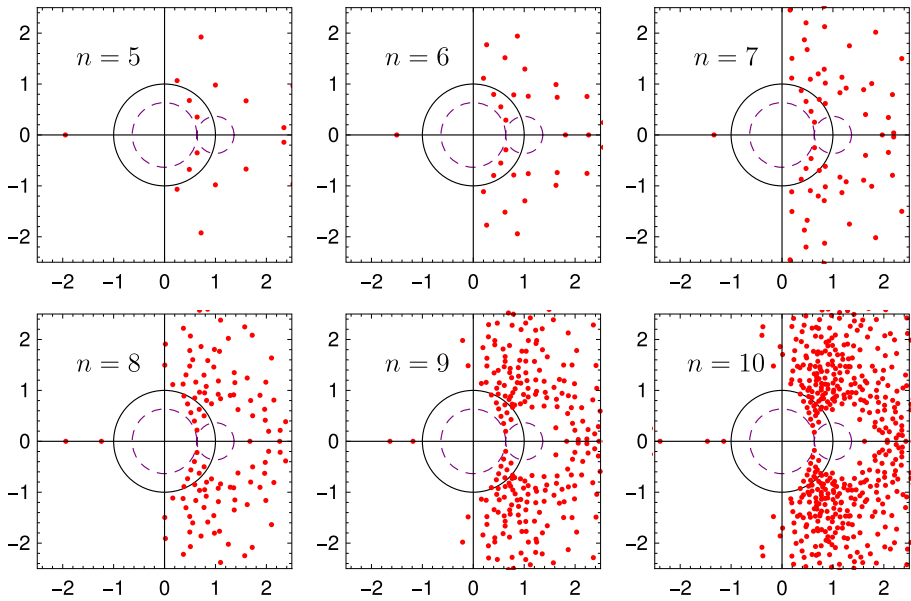
Fig. 6 Location of the poles of $S_{[n]}$ **as a function of** p in the complex plane for different n. The black circle is the complex unit circle and the dashed circles have radius p_c around $p = 0$ and $1 - p_c$ around $p = 1$

in complete agreement with [10], and that the critical exponent for $S_\mathbb{Z}(q) \overset{q \uparrow q_c}{\sim} (q_c - q)^\beta$ is $\beta \approx 0.277$, which suggests that it is in the directed-percolation (DP) universality class alongside several variants of the contact process [9,14,25].

3.3 Reaching One End of the Chain from the Other

Another quantity we considered in Sect. 1.1 is the probability of ever activating one end point of a finite chain, when we start the process with only a single active vertex at the other end. Let us consider the length-n chain, and suppose we start the DBS process with a single active vertex at site 1. As in Eq. (2), we consider

$$S_{[n]}(p) = \mathbb{P}(\mathrm{BA}^{(\{n\})} \mid \mathrm{start} \{1\}).$$

Note that in order to satisfy property (i) of the PLUP definition, the initial state needs to be $\{1\}$ with probability p and $\emptyset$ with probability $1 - p$. To get the above definition of $S_{[n]}(p)$ with a deterministic starting state one can then simply divide by p. The power-series coefficients of $S_{[n]}(p)$ stabilize, which follows from Lemma 5 by letting $X = \{n\}$ and $Y = \{1\}$. However, as suggested by Fig. 1, the limiting power series around $p = 0$ will become the zero function and it is therefore not so interesting. Instead, we can take the power series centered around $p = 1$ and it turns out that also there the coefficients stabilize. We prove this below. Define $q = 1 - p$.

Similarly to what we did for $R_{(n)}(p)$ we can write $S_{[n]}(q)$ using a matrix inverse. We will start the process in the (deterministic) state with a single active vertex at location 1, denoted by the probability vector $\delta_{\{1\}}$. Define $\mathcal{A}_n = \{A \subseteq [n] \mid n \in A\}$, the set of all states where vertex n is active. Let $M_{[n]}$ be the transition matrix for the DBS process on the chain of length n. Define the matrix $\tilde{M}_{[n]}$ as $M_{[n]}$ but with some entries set to zero. Set the row and column

of the all-inactive state $\emptyset$ to zero, $(\tilde{M}_{[n]})_{A,\emptyset} = (\tilde{M}_{[n]})_{\emptyset,A} = 0$ for all $A \subseteq [n]$. Furthermore set all rows $A \in \mathcal{A}_n$ to zero: $(\tilde{M}_{[n]})_{A,A'} = 0$ for all $A' \subseteq [n]$. That is, whenever vertex n is active there is no outgoing transition. Denote by $\chi_{\mathcal{A}_n}$ the vector that is 1 for all $A \in \mathcal{A}_n$ and zero everywhere else. We have

$$
\begin{aligned}
S_{[n]}(q) &= \mathbb{P}(\text{vertex } n \text{ becomes active}) \\
&= \sum_{k \geq 0} \mathbb{P}(\text{vertex } n \text{ activates for the first time at update } k) \\
&= \sum_{k \geq 0} \delta_{\{1\}} \cdot \tilde{M}_{[n]}^k \cdot \chi_{\mathcal{A}_n} \\
&= \delta_{\{1\}} \cdot (\text{Id} - \tilde{M}_{[n]})^{-1} \cdot \chi_{\mathcal{A}_n} \quad \text{(by the geometric series)} \\
&= \sum_{k \geq 0} b_k^{[n]} q^k
\end{aligned}
\tag{10}
$$

With the same argument as before we see that $S_{[n]}$ must be a fraction of two polynomials in p (and also in q). The poles of $S_{[n]}$ are shown in Fig. 6 where $S_{[n]}$ is considered a function of p to be comparable with $R_{(n)}(p)$. The coefficients $b_k^{[n]}$ of the q power series are shown in Table 2.

Lemma 6 *The coefficients $b_k^{[n]}$ of the power series of $S_{[n]}(q)$ in Eq. (10) stabilize.*

Proof Let $\text{RI}^{(\{n\})}$ and its complement $\text{BA}^{(\{n\})}$ be as defined in Definition 2. In the following we assume that the starting state is $\{1\}$ with probability p and $\emptyset$ with probability $1 - p$, so the process is a PLUP. We have $S_{[n]}(p) = \frac{1}{p} \cdot \mathbb{P}(\text{BA}^{(n)})$, since $S_{[n]}(p)$ has a deterministic starting state. By Lemma 3 we have $\mathbb{P}_{[n]}(\text{RI}^{(\{n-1\})}) = \mathbb{P}_{[n-1]}(\text{RI}^{(\{n-1\})})$. Consider $1 - p S_{[n]}$, i.e. the probability that the n-th vertex is *not* activated. We have

$$
\begin{aligned}
1 - p S_{[n]} &= \mathbb{P}_{[n]}(\text{RI}^{(\{n\})}) && \text{(definition of } S_{[n]}) \\
&= \mathbb{P}_{[n]}(\text{RI}^{(\{n-1\})} \cap \text{RI}^{(\{n\})}) + \mathbb{P}_{[n]}(\text{BA}^{(\{n-1\})} \cap \text{RI}^{(\{n\})}) && \text{(partition of events)} \\
&= \mathbb{P}_{[n-1]}(\text{RI}^{(\{n-1\})}) + \mathbb{P}_{[n]}(\text{BA}^{(\{n-1\})} \cap \text{RI}^{(\{n\})}) && \text{(Lemma 3)} \\
&= 1 - p S_{[n-1]} + \mathbb{P}_{[n]}(\text{BA}^{(\{n-1\})} \cap \text{RI}^{(\{n\})}).
\end{aligned}
$$

Note that for the event $(\text{BA}^{(\{n-1\})} \cap \text{RI}^{(\{n\})})$ to hold, all vertices $1, \dots, n - 1$ must have been active. Since the process terminates with probability 1, this means all those vertices must also have been deactivated at least once. In the DBS process a deactivation is $O(q)$, so every terminating path of the Markov Chain that is in this set has a factor of at least q^{n-1} associated to it, hence $\mathbb{P}_{[n]}(\text{BA}^{(\{n-1\})} \cap \text{RI}^{(\{n\})}) = O(q^{n-1})$. Here we use the absolute convergence of certain power series in q, which we prove in Lemma 10 in Appendix C. We see that $S_{[n]}(q) - S_{[n-1]}(q) = O(q^{n-1})$ so the coefficients stabilize. $\square$

A Decay of Correlations

Recall Corollary 1.

Corollary 1 (Decay of correlations) *Let M_G be a parametrized local-update process on the graph $G = (V, E)$. If $X, Y \subseteq V$ and L_X, L_Y are local events on X and Y respectively, then*

$$\text{Cov}(L_X, L_Y) = \mathbb{P}_G(L_X \cap L_Y) - \mathbb{P}_G(L_X)\mathbb{P}_G(L_Y) = O(p^{d(X,Y)-1}), \tag{5}$$

and

$$\mathbb{P}_G(\text{BA}^{(X)} \cap \text{BA}^{(Y)}) - \mathbb{P}_G(\text{BA}^{(X)})\mathbb{P}_G(\text{BA}^{(Y)}) = O(p^{d(X,Y)+1}). \tag{6}$$

Proof First observe that if $d(X, Y) = \infty$, it means that either X and Y are in different connected components of G, or one of them is the empty set, therefore L_X and L_Y are independent events, so the statement holds.

Note that due to Property (i) the only path which has a non-zero constant term is the trivial path, when every vertex is initialized as inactive, thus the constant term of the probability of any local event is either 0 or 1. Also the constant term of $\mathbb{P}_G(L_X \cap L_Y)$ is 1 if and only if the constant terms of both $\mathbb{P}_G(L_X)$ and $\mathbb{P}_G(L_Y)$ are 1, which concludes the $d(X, Y) = 0$ case.

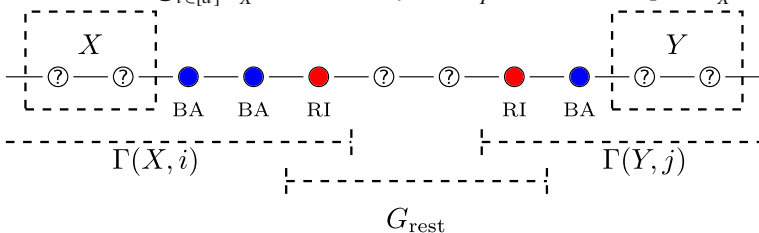Note that by De Morgan's law, (6) is equivalent with

$$\mathbb{P}_G(\text{RI}^{(X)} \cap \text{RI}^{(Y)}) - \mathbb{P}_G(\text{RI}^{(X)})\mathbb{P}_G(\text{RI}^{(Y)}) = O(p^{d(X,Y)+1}). \tag{11}$$

Now we proceed by induction on $d(X, Y)$. Assume (5)–(6) hold for $d(X, Y) = d - 1$. We will prove the statement for $d(X, Y) = d$. We apply a similar idea as in the proof of Lemma 5. Define

$$L_X^i := L_X \cap \text{RI}^{(\bar{\partial}(X,i))} \cap \bigcap_{j \in [i-1]} \text{BA}^{(\bar{\partial}(X,j))} \quad \text{for } i \in [d-1],$$

$$L_X^d := L_X \cap \bigcap_{j \in [d-1]} \text{BA}^{(\bar{\partial}(X,j))}.$$

When L_X^i holds, everything at distance i remains inactive, but for all distances $j \leq i - 1$ there exist vertices that become active at that distance. These events form a partition $L_X = \dot{\bigcup}_{i \in [d]} L_X^i$, and similarly for L_Y^i. Below we depict $L_X^i \cap L_Y^j$ graphically.

We will show the inductive step for both (5)–(6) at the same time, for which we introduce a number c such that $c = 1$ if $L_X = \text{BA}^{(X)}$ and $L_Y = \text{BA}^{(Y)}$, and $c = -1$ otherwise. By Lemma 4

$$\mathbb{P}(L_X^i \cap L_Y^j) = O(p^{i+j-1+c}) \quad \text{and} \quad \mathbb{P}(L_X^i) \cdot \mathbb{P}(L_Y^j)$$
$$= O(p^{i+j-1+c}), \tag{12}$$

for any graph on which the events are defined. Since the events form a partition, we have

$$\mathbb{P}_G(L_X \cap L_Y) = \sum_{i,j \in [d]} \mathbb{P}_G(L_X^i \cap L_Y^j) \quad \text{and} \quad \mathbb{P}_G(L_X) \cdot \mathbb{P}_G(L_Y)$$

$$= \sum_{i,j \in [d]} \mathbb{P}_G(L_X^i) \cdot \mathbb{P}_G(L_Y^j),$$

so it is sufficient to prove the statement for each i, j separately, i.e. we want to show

$$\mathbb{P}_G(L_X^i \cap L_Y^j) - \mathbb{P}_G(L_X^i)\mathbb{P}_G(L_Y^j) = O(p^{d+c}).$$

When $i + j - 1 \geq d$ then it is trivial by (12). Now fix i, j such that $i + j \leq d$ and define $G_{\text{rest}}^{i,j} := G \setminus (\Gamma(X, i-1) \cup \Gamma(Y, i-1))$, as indicated in the diagram. The $\text{RI}^{(\cdot \cdot)}$ events split the graph in three parts, so we have

$$
\begin{aligned}
\mathbb{P}_G(L_X^i \cap L_Y^j) &= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \Pi^{(\overline{\partial}(X,i))}) \\
&\quad \cdot \mathbb{P}_{\Gamma(Y,j)}(L_Y^j \mid \Pi^{(\overline{\partial}(Y,j))}) \cdot \mathbb{P}_{G_{\text{rest}}^{i,j}}(\text{RI}^{(\overline{\partial}(X,i))} \cap \text{RI}^{(\overline{\partial}(Y,j))}) && \text{(using Lemma 3 twice)} \\
&= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \Pi^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{\Gamma(Y,j)}(L_Y^j \mid \Pi^{(\overline{\partial}(Y,j))}) \\
&\quad \cdot \left[\mathbb{P}_{G_{\text{rest}}^{i,j}}(\text{RI}^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{G_{\text{rest}}^{i,j}}(\text{RI}^{(\overline{\partial}(Y,j))}) + O(p^{(d-i-j)+1}) \right] && \text{(by induction of (11))} \\
&= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \Pi^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{\Gamma(Y,j)}(L_Y^j \mid \Pi^{(\overline{\partial}(Y,j))}) \\
&\quad \cdot \mathbb{P}_{G_{\text{rest}}^{i,j}}(\text{RI}^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{G_{\text{rest}}^{i,j}}(\text{RI}^{(\overline{\partial}(Y,j))}) + O(p^{d+c}) && \text{(by (12))} \\
&= \mathbb{P}_{\Gamma(X,i)}(L_X^i \mid \Pi^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{\Gamma(Y,j)}(L_Y^j \mid \Pi^{(\overline{\partial}(Y,j))}) \\
&\quad \cdot \mathbb{P}_{G \setminus \Gamma(X,i-1)}(\text{RI}^{(\overline{\partial}(X,i))}) \cdot \mathbb{P}_{G \setminus \Gamma(Y,j-1)}(\text{RI}^{(\overline{\partial}(Y,j))}) + O(p^{d+c}) && \text{(by Lemma 5 and (12))} \\
&= \mathbb{P}_G(L_X^i) \cdot \mathbb{P}_G(L_Y^j) + O(p^{d+c}) && \text{(using Lemma 3 twice)}
\end{aligned}
$$

$\square$

B Absolute Convergence

Recall Lemma 4.

Lemma 4 *Let M_G be a parametrized local-update process on the graph G with vertex set V. Let $\{X_1, \ldots, X_k\}$ be a collection of disjoint vertex subsets $X_i \subseteq V$ and let E be an event. If $E \subseteq \bigcap_i \text{BA}^{(X_i)}$, then $\mathbb{P}(E) = O(p^k)$. Furthermore if $S \subseteq V$ then also $\mathbb{P}(E \mid \Pi^{(S)}) = O(p^k)$.*

When the event E holds, each set X_i contains a vertex that becomes active, and by PLUP properties (i) and (iii) any activation (either during initialization or later) is $O(p)$. Therefore, for any path ξ of the Markov Chain with $\xi \in E$ we have $\mathbb{P}(\xi) = O(p^k)$, where $\mathbb{P}(\xi)$ is a polynomial in p. We have $\mathbb{P}(E) = \sum_{\xi \in E} \mathbb{P}(\xi)$ by definition. This is a sum over infinitely many polynomials, and by considering $\mathbb{P}(E)$ as a power series in p we are effectively regrouping terms in this sum. In this section we prove the absolute convergence of certain series that allows for this regrouping. We start with some notation.

Definition 7 *(Paths)* Define a *path* of length k as an initialization and sequence of k updates, where we only count steps in which an active vertex was selected. We write a path ξ as

$$\xi = ((\text{initialize to } A_0), (v_1, R_1), (v_2, R_2), \ldots, (v_k, R_k)).$$

Here v_i denotes the vertex that was selected in the i-th step and $R_i \subseteq \Gamma(v_i)$ is the result of the corresponding update that happened afterwards. After t steps, the state of the process is $A_t = (A_{t-1} \setminus \Gamma(v_t)) \cup R_t$. We say a path is terminating if $A_k = \emptyset$. Denote by PATHS$_{A,k}$ the set of all paths ξ that initialize to A and have length k.

For a general PLUP we have $\mathbb{P}(\xi) = \mathbb{P}(A_0)\mathbb{P}((v_1, R_1) \mid A_0)\mathbb{P}((v_2, R_2) \mid A_1) \cdots \mathbb{P}((v_k, R_k) \mid A_{k-1})$ where the polynomial $\mathbb{P}(A_0)$ is the probability of starting in state A_0 and $\mathbb{P}((v_t, R_t) \mid A_{t-1})$ are polynomials satisfying property (iii) of the PLUP definition. For the DBS process on the cycle these polynomials take the specific form $\mathbb{P}(\xi) = \mathbb{P}(A_0)Z_\xi \, p^{|R_1|+\cdots+|R_k|}(1 - p)^{3k-|R_1|+\cdots+|R_k|}$ where Z_ξ is some p-independent factor.

Definition 8 *(Polynomials)* Let $Q(p) = a_m p^m + a_{m+1} p^{m+1} + \cdots + a_M p^M$ be a polynomial where $a_m \neq 0$ and $a_M \neq 0$. Define mindeg$(Q(p)) = m$, maxdeg$(Q(p)) = M$ and define by $\|Q\|_{\text{abs}}$ the polynomial obtained by taking the absolute values of the coefficients:

$$\|Q\|_{\text{abs}}(p) = |a_m|p^m + |a_{m+1}|p^{m+1} + \cdots + |a_M|p^M.$$

By the triangle inequality $\|f \cdot g\|_{\text{abs}}(p) \leq \|f\|_{\text{abs}}(p) \cdot \|g\|_{\text{abs}}(p)$ for any polynomials f, g, and $p \geq 0$.

Lemma 7 *The polynomials $\mathbb{P}(\text{PATHS}_{A,k})$ and $\mathbb{P}(\xi)$ for any $\xi \in \text{PATHS}_{A,k}$ all satisfy*

$$\text{mindeg}(\cdot) \geq c \cdot (k - |A|) + \text{mindeg}(\mathbb{P}(A)) , \qquad \text{maxdeg}(\cdot) \leq c' \cdot k + \text{maxdeg}(\mathbb{P}(A)).$$

Here $0 < c < c'$ are constants depending on the particular process and $\mathbb{P}(A)$ is the probability of starting in state A (a polynomial).

Proof Note that

$$\mathbb{P}(\text{PATHS}_{A,k}) = \sum_{\xi \in \text{PATHS}_{A,k}} \mathbb{P}(\xi)$$

is a sum over finitely many polynomials. It is sufficient to prove the statement for each ξ and it then follows for the sum. Let ξ be a path as described in Definition 7. As stated in the text below Definition 7 we have $\mathbb{P}(\xi) = \mathbb{P}(A_0)\mathbb{P}((v_1, R_1) \mid A_0)\mathbb{P}((v_2, R_2) \mid A_1) \cdots \mathbb{P}((v_k, R_k) \mid A_{k-1})$ where $A_t \subseteq [n]$ is the state after t steps and $A_0 = A$. Let c' be the degree of the highest order term of any possible local-update step of this process (finitely many possibilities) then maxdeg$(\mathbb{P}(\xi)) \leq c' \cdot k + \text{maxdeg}(\mathbb{P}(A))$.

Note mindeg$(\mathbb{P}(\xi)) = $ mindeg$(\mathbb{P}(A)) + $ mindeg$(\mathbb{P}((v_1, R_1) \mid A_0)) + \cdots + $ mindeg $(\mathbb{P}((v_k, R_k) \mid A_{k-1}))$. If $|A_t| - |A_{t-1}| \geq 0$ then either $A_t = A_{t-1}$ or $|A_t \setminus A_{t-1}| > 0$. By property (iii) of the PLUP definition we therefore have that $|A_t| - |A_{t-1}| \geq 0$ implies mindeg$(\mathbb{P}((v_t, R_t) \mid A_{t-1})) \geq 1$. Furthermore, $|A_t| - |A_{t-1}| \leq d_{\max}$ where $d_{\max}$ is the maximum degree of the vertices in G. Therefore we have

$$\text{mindeg}(\mathbb{P}((v_t, R_t) \mid A_{t-1})) \geq \frac{1}{d_{\max} + 1}(1 + |A_t| - |A_{t-1}|)$$

Summing both sides over t we obtain mindeg$(\mathbb{P}(\xi)) - $ mindeg$(\mathbb{P}(A)) \geq \frac{1}{d_{\max}+1}(k + |A_k| - |A|)$. This proves the lemma with $c = \frac{1}{d_{\max}+1}$. $\qquad \square$

Lemma 8 *There is a constant $\delta > 0$ such that, for any polynomial $f(k)$, the following series is absolutely convergent for $p \in [0, \delta]$:*

$$\sum_{k=0}^{\infty} \sum_{A \subseteq [n]} \sum_{\xi \in \text{PATHS}_{A,k}} f(k) \, \|\mathbb{P}(\xi)\|_{\text{abs}} < \infty.$$

Proof We have $\mathbb{P}(\xi) = \mathbb{P}(A_0)\mathbb{P}((v_1, R_1) \mid A_0)\mathbb{P}((v_2, R_2) \mid A_1) \cdots \mathbb{P}((v_k, R_k) \mid A_{k-1})$. The polynomials $P_t := \mathbb{P}((v_t, R_t) \mid A_{t-1})$ come from a finite set of polynomials: for each vertex v there are at most $2^{|\Gamma(v)|}$ possible updates and there are at most n vertices. Therefore there is a constant C such that for all these polynomials

$$\|P_t\|_{\mathrm{abs}}(p) \leq C\, p^{\mathrm{mindeg}(P_t)}, \quad \forall p \in [0, 1].$$

By Lemma 7 there is a c such that

$$\|\mathbb{P}(\xi)\|_{\mathrm{abs}} \leq \|\mathbb{P}(A_0)\|_{\mathrm{abs}} \|P_1\|_{\mathrm{abs}} \cdots \|P_k\|_{\mathrm{abs}} \leq \|\mathbb{P}(A_0)\|_{\mathrm{abs}} C^k\, p^{\mathrm{mindeg}(P_1) + \cdots \mathrm{mindeg}(P_k)}$$

$$\leq \|\mathbb{P}(A)\|_{\mathrm{abs}} C^k p^{c \cdot (k - |A|)}.$$

There are at most $(2^{d_{\max}} n)^k$ paths of length k for a fixed starting state so we have

$$\sum_{k=0}^{\infty} \sum_{A \subseteq [n]} \sum_{\xi \in \mathrm{PATHS}_{A,k}} f(k)\, \|\mathbb{P}(\xi)\|_{\mathrm{abs}} \leq \sum_{k=0}^{\infty} \sum_{A \subseteq [n]} f(k)\, \|\mathbb{P}(A)\|_{\mathrm{abs}} (2^{d_{\max}} n)^k C^k p^{c(k - |A|)}$$

Since there are finitely many (2^n) starting states A, the whole expression is absolutely convergent for $p < (2^{d_{\max}} nC)^{-1/c}$, since f is a polynomial. □

Denote by $\mathrm{TPATHS}_{A,k}$ the set of all *terminating* paths that initialize to A and have length k. By the above lemma, the process terminates with probability 1 for small enough p, i.e. for $p \in [0, \delta]$

$$\sum_{k=0}^{\infty} \sum_{A \subseteq [n]} \sum_{\xi \in \mathrm{TPATHS}_{A,k}} \mathbb{P}(\xi) = 1.$$

This also implies that, up to measure zero events, any local event E is a subset of the set of all terminating paths. Therefore the powerseries $\mathbb{P}(E) = \sum_{\xi \in E} \mathbb{P}(\xi)$ is absolutely convergent and we are allowed to rearrange the polynomials in this sum. We can now finish the proof of Lemma 4.

Proof For all $\xi \in E$ we have $\mathbb{P}(\xi) = O(p^k)$. For $p \in [0, \delta]$ we have $\mathbb{P}(E) = \sum_{j=k}^{\infty} a_j p^j$ by Lemma 8. By uniqueness of power series, this equality holds for all p up to the radius of convergence. We conclude $\mathbb{P}(E) = O(p^k)$. When the process is conditioned on $\Pi^{(S)}$ then it is simply a new PLUP so the same proof holds. □

For the DBS process, in the context of Sect. 3.2, we can slightly refine Lemma 7.

Lemma 9 *Let $\rho^{(0)}$, $M_{(n)}$ and J be as defined in Sect. 3.2. The polynomial $\rho^{(0)} \cdot M_{(n)}^k \cdot J^T = \sum_{A \subseteq [n]} \mathbb{P}(\mathrm{PATHS}_{A,k})$ in p has lowest-order term at least p^k and highest-order term at most p^{n+3k}.*

Proof We repeat the proof of Lemma 7. Note that $\mathrm{mindeg}(\mathbb{P}((v_t, R_t) \mid A_{t-1})) \geq 1 + |A_t| - |A_{t-1}|$ for the DBS process, so $c = 1$. For DBS, $c' = 3$ which is the maximum degree of the local update rule (p^3 occurs when all three resampled vertices become active). The lemma then follows by noting that $\mathbb{P}(A) = p^{|A|}(1 - p)^{n-|A|}$ in the starting state $\rho^{(0)}$. □

This lemma is convenient for the computation of the $R_{(n)}(p)$ power series. It implies that the term p^j is only present in those polynomials $\rho^{(0)} \cdot M_{(n)}^k \cdot J^T$ for which $\lceil \frac{j-n}{3} \rceil \leq k \leq j$. To compute the power-series coefficient $a_j^{(n)}$ it is sufficient to consider this finite set of polynomials. In other words, in order to compute $R_{(n)}(p)$ up to k-th order in p, it suffices to consider only the first k steps of the DBS process. We use this observation to compute the coefficients of the $n \geq 18$ series by computing matrix powers symbolically in p, see Table 1.

C Absolute Convergence for the *q* Power Series

We now turn our attention to the $S_{[n]}(q)$ series defined in Eq. (10). This process starts with a single active vertex at position 1, i.e. $A = \{1\}$, and we look at the probability that vertex n is never activated, $\mathbb{P}(\text{RI}^{(\{n\})} \mid \text{start } A)$, as a function of $q = 1 - p$. To prove the absolute convergence of such series for general PLUPs we introduce some additional assumptions. We now consider the update polynomials as a function of $q = 1 - p$. The update rule for a single time step should satisfy the following additional two properties.

- For $q = 0$ the probability that an active vertex becomes inactive is zero. This implies that any deactivation has probability $O(q)$.
- There is a $c > 0$ such that if $q = 0$, then for any inactive vertex v that has an active neighbor, the probability of activating v is at least c.

These properties are satisfied by the CP and DBS processes. Note that c is independent of q but will generally depend on the system size n.

Lemma 10 *Consider a PLUP that satisfies the above two properties. Let $X \subset V$ be any subset of vertices such that its boundary $B = \bar{\partial}(X; 1)$ is not empty. Let the starting state be $A \subseteq X$. Define $\text{RI}_k^{(B)}$ as the set of all lenght-k paths for which all vertices in B remained inactive. Then the following series converges for small enough q*

$$\sum_{k \geq 0} \sum_{\xi \in \text{RI}_k^{(B)}} \|\mathbb{P}(\xi)\|_{\text{abs}}(q) < \infty.$$

Moreover, $\text{RI}^{(B)} \subset \bigcup_{k \geq 0} \text{RI}_k^{(B)}$, up to zero-probability events, so we can regroup terms in the series $\mathbb{P}(\text{RI}^{(B)})$.

Proof In each state A, the process can do at most a finite amount ($2^{d_{\max}}n$, where $d_{\max}$ is the maximum degree of the graph) of possible transitions, including both selection and update dynamics. Let $Q_j^{(A)}(q)$ be the j-th non-zero transition polynomial in state A (now including selection dynamics) so that $\sum_j Q_j^{(A)}(q) = 1$. This holds in particular for $q = 0$ so the constant terms of $Q_j^{(A)}(q)$ are non-negative and sum to 1. Hence, there is a $\delta > 0$ and another constant such that for all state A, and $q \in [0, \delta]$ we have $Z^{(A)}(q) := \sum_j \left\| Q_j^{(A)} \right\|_{\text{abs}}(q) \leq 1 + \text{const}\cdot q$. Define new normalized functions $\tilde{Q}_j^{(A)}(q) := \frac{1}{Z^{(A)}(q)} \left\| Q_j^{(A)} \right\|_{\text{abs}}(q)$ and consider the same process but with the transition polynomials $Q_j^{(A)}(q)$ replaced by the rational functions $\tilde{Q}_j^{(A)}(q)$. We will denote probabilities for this process by $\tilde{\mathbb{P}}$.

For any path ξ of length k we now have $\|\mathbb{P}(\xi)\|_{\text{abs}} = \tilde{\mathbb{P}}(\xi) \prod_{j=0}^{k-1} Z^{(A_j)}$, where A_j is the state after the j-th transition. This allows us to bound the sum for $q \in [0, \delta]$ as follows

$$\sum_{\xi \in \text{RI}_k^{(B)}} \|\mathbb{P}(\xi)\|_{\text{abs}}(q) \leq (1 + \text{const} \cdot q)^k \, \tilde{\mathbb{P}}\left(\text{RI}_k^{(B)}\right)(q).$$

We proceed by bounding $\tilde{\mathbb{P}}(\text{RI}_k^{(B)})$. Define the following random variables. Let $I_t \in \{0, 1\}$ be 1 if any active vertex got inactivated in step t. Let $G_t \in \{0, 1\}$ be 1 if any inactive vertex got activated in step t (G stands for grow). Let $I = \sum_{t=1}^k I_t$ and $G = \sum_{t=1}^k G_t$. We always have $G \leq |X| + (1 + d_{\max})I$, because after $|X|$ activations any other activation requires a

deactivation first, and in a single step the process can deactivate at most $1 + d_{max}$ vertices at once.

By the second additional property there is a c such that any inactive neighbor of an active vertex can be activated with probability at least c when $q = 0$, i.e., if there are inactive vertices in step t then $\mathbb{P}(G_t = 1) \geq c$. The tilde process coincides with the regular one for $q = 0$, and by continuity there is a $q_0 \in (0, \delta]$ such that for all $q \in [0, q_0]$ we have $\tilde{\mathbb{P}}(G_t = 1) \geq c/2$.

Claim 1 For all $q \in [0, q_0]$, the random variable G satisfies

$$\tilde{\mathbb{P}}\left(\mathrm{RI}_k^{(B)} \cap \left(G \leq \frac{c}{4}k\right)\right)(q) \leq \exp(-ck/16).$$

Proof Since $\mathrm{RI}_k^{(B)}$ holds there is always at least one inactive vertex with an active neighbor. We have $\tilde{\mathbb{P}}(G_t = 1) \geq c/2$. Define k i.i.d. Bernoulli variables C_t with success probability $c/2$ and $C = \sum_{t=1}^{k} C_t$. The expectation of C is $\mathbb{E}(C) = \frac{c}{2}k$ and using the Chernoff bound we can bound the probability that C deviates far from its mean:

$$\mathbb{P}\left(C \leq \frac{c}{4}k\right) \leq e^{-ck/16}$$

We use a coupling argument to compare the G_t variables with the C_t's. Let U_t be i.i.d. uniform $[0, 1]$ variables. Define C_t to be 1 if $U_t < c/2$ so the C_t's are indeed i.i.d. Bernoulli variables with the correct distribution.

For G_t run the process, and in each step first compute the true probability $p_t = \tilde{\mathbb{P}}(G_t = 1 \mid \text{history})$, so $p_t \geq c/2$. Now use the randomness of U_t to decide what happens, i.e. define $G_t = 1$ if and only if $U_t < p_t$. Then continue the process conditioned on the value of G_t. This way, the G_t variables come from the correct distribution but they are coupled to the C_t's. We see $C_t = 1$ implies $G_t = 1$ so $G \geq C$ and therefore $\mathbb{P}(G \leq \frac{c}{4}k) \leq \mathbb{P}(C \leq \frac{c}{4}k) \leq e^{-ck/16}$. □

Now we continue the proof of Lemma 10. We partition the $\mathrm{RI}_k^{(B)}$ event as

$$\tilde{\mathbb{P}}(\mathrm{RI}_k^{(B)}) = \tilde{\mathbb{P}}\left(\mathrm{RI}_k^{(B)} \cap \left(G \leq \frac{c}{4}k\right)\right) + \tilde{\mathbb{P}}\left(\mathrm{RI}_k^{(B)} \cap \left(G > \frac{c}{4}k\right)\right).$$

The first term is bounded by the claim above. From $G \leq |X| + (1 + d_{max})I$ it follows that $I \geq \frac{G - |X|}{1 + d_{max}}$. By the first property every deactivation has an update step that is $O(q)$, in the original process. For the corresponding transition polynomials there is a constant such that $Q_j^{(A)}(q) \leq \mathrm{const}' \cdot q$ for all $q \in [0, 1]$, which implies that there is another constant such that $\tilde{Q}_j^{(A)}(q) \leq \mathrm{const}'' \cdot q$ for small enough q. Therefore, there is a constant such that for a single step of the tilde process we have $\tilde{\mathbb{P}}(\text{deactivation}) \leq \mathrm{const}''' \cdot q$. The probability of I deactivations is therefore at most $(\mathrm{const}''' \cdot q)^I$. The second term is therefore bounded by

$$\tilde{\mathbb{P}}\left(\mathrm{RI}_k^{(B)} \cap \left(G > \frac{c}{4}k\right)\right)(q) \leq (\mathrm{const}''' \cdot q)^{\frac{1}{1+d_{max}}(\frac{c}{4}k - |X|)}.$$

We see that for small enough q

$$\sum_{k \geq 0} \sum_{\xi \in \mathrm{RI}_k^{(B)}} \|\mathbb{P}(\xi)\|_{abs}(q) \leq \sum_{k \geq 0} (1 + \mathrm{const} \cdot q)^k \, \tilde{\mathbb{P}}(\mathrm{RI}_k^{(B)})(q)$$

$$\leq \sum_{k \geq 0} (1 + \mathrm{const} \cdot q)^k \left(\exp\left(-\frac{c}{16}k\right) + (\mathrm{const}''' \cdot q)^{\frac{1}{1+d_{max}}(\frac{c}{4}k - |X|)}\right)$$

is convergent. □

D Proving that $a_k^{(k+1)} = a_k^{(n)}$ for all $n > k$

In the main text we looked at $R_{(n)}(p)$ and saw that the coefficients of this series stabilize. Our main theorem, however, only shows that this stabilization happens for $n > 2k$ since any new vertex added by going from n to $n+1$ is at distance at most $n/2$ from all existing vertices. In this section we prove the stronger statement, namely that the coefficients stabilize for $n > k$. Let

$$P_C := \mathrm{RI}^{(\bar{\partial}C)} \cap \bigcap_{v \in C} \mathrm{BA}^{(\{v\})}$$

be the event that every vertex in C becomes active, and the boundary remains inactive.

The intuition of the following theorem is similar to that of Corollary 1. A site can only realize the length of the cycle after an interaction chain was formed around the cycle, implying that every vertex was activated at least once.

Theorem 2 $R_{(n)} = \mathbb{E}_{[-m,m]}(\#\mathrm{ASEL}(0)) + O(p^n)$ for all $m \geq n \geq 3$, thus $R_{(n)} - R_{(m)} = O(p^n)$.

Proof In the proof we identify the sites of the n-cycle with the $\mod n$ remainder classes. We have $R_{(n)}(p) = \mathbb{E}_{(n)}(\#\mathrm{ASEL}(0))$ by translation invariance, and this expectation is equal to $\sum_{k=1}^{\infty} \mathbb{P}_{(n)}(\#\mathrm{ASEL}(0) \geq k)$. Let us abbreviate the event as $X = (\#\mathrm{ASEL}(0) \geq k)$.

$$\mathbb{P}_{(n)}(X) = \sum_{\substack{v,w \in [n] \\ v+w \leq n+1}} \mathbb{P}_{(n)}(X \cap \underbrace{P_{[-v+1,w-1]}}_{P_{v,w} :=}) \qquad \text{(partition)}$$

$$= \sum_{\substack{v,w \in [n] \\ v+w \leq n}} \mathbb{P}_{(n)}(X \cap P_{v,w}) + O(p^n)$$

$$= \sum_{\substack{v,w \in [n] \\ v+w \leq n}} \mathbb{P}_{[-v,w]}(X \cap P_{v,w} \mid \Pi^{(\{-v,w\})})\mathbb{P}_{[w,n-v]}(\mathrm{RI}^{(\{w,n-v\})}) + O(p^n) \quad \text{(by Lemma 3)}$$

$$= \sum_{\substack{v,w \in [n] \\ v+w \leq n}} \mathbb{P}_{[-v,w]}(X \cap P_{v,w} \mid \Pi^{(\{-v,w\})})$$
$$\cdot \left[\left(\mathbb{P}_{[w,n-v]}(\mathrm{RI}^{(w)})\right)^2 + O(p^{n-v-w+1})\right] + O(p^n) \qquad \text{(Corollary 1 and Equation (11)}$$

$$= \sum_{\substack{v,w \in [n] \\ v+w \leq n}} \mathbb{P}_{[-v,w]}(X \cap P_{v,w} \mid \Pi^{(\{-v,w\})})$$
$$\cdot \left[\mathbb{P}_{[-m,-v]}(\mathrm{RI}^{(-v)})\mathbb{P}_{[w,m]}(\mathrm{RI}^{(w)}) + O(p^{n-v-w+1})\right] + O(p^n) \qquad \text{(by Lemma 5)}$$

$$= \sum_{\substack{v,w \in [n] \\ v+w \leq n}} \mathbb{P}_{[-v,w]}(X \cap P_{v,w} \mid \Pi^{(\{-v,w\})})\mathbb{P}_{[-m,-v]}(\mathrm{RI}^{(-v)})\mathbb{P}_{[w,m]}(\mathrm{RI}^{(w)}) + O(p^n)$$
$$\text{(since } |P_{v,w}| = v + w - 1)$$

$$= \sum_{\substack{v,w \in [n] \\ v+w \leq n}} \mathbb{P}_{[-m,m]}(X \cap P_{v,w}) + O(p^n) \qquad \text{(by Lemma 3)}$$

$$= \mathbb{P}_{[-m,m]}(X) + O(p^n) \qquad \text{(partition)}$$

We conclude the proof by observing

$$\sum_{k=1}^{\infty} \mathbb{P}_{[-m,m]}(\#\mathrm{ASEL}(0) \geq k) + O(p^n) = \mathbb{E}_{[-m,m]}(\#\mathrm{ASEL}(0)) + O(p^n).$$

References

1. Bak, P.: How Nature Works. Copernicus, New York (1996). https://doi.org/10.1007/978-1-4757-5426-1. The science of self-organized criticality
2. Bak, P., Sneppen, K.: Punctuated equilibrium and criticality in a simple model of evolution. Phys. Rev. Lett. **71**, 4083–4086 (1993). https://doi.org/10.1103/PhysRevLett.71.4083
3. Baker, G.A., Hunter, D.L.: Methods of series analysis. II. Generalized and extended methods with application to the ising model. Phys. Rev. B **7**, 3377–3392 (1973). https://doi.org/10.1103/PhysRevB.7.3377
4. Bandt, C.: The discrete evolution model of Bak and Sneppen is conjugate to the classical contact process. J. Stat. Phys. **120**(3), 685–693 (2005). https://doi.org/10.1007/s10955-005-5965-x
5. Barbay, J., Kenyon, C.: On the discrete Bak–Sneppen model of self-organized criticality. In: Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '01, pp. 928–933. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2001). http://dl.acm.org/citation.cfm?id=365411.365814
6. Catarata, J.D., Corbett, S., Stern, H., Szegedy, M., Vyskocil, T., Zhang, Z.: The Moser-Tardos resample algorithm: where is the limit? (an experimental inquiry). In: Fekete, S.P., Ramachandran, V. (eds.) Proceedings of the Ninteenth Workshop on Algorithm Engineering and Experiments, ALENEX 2017, Barcelona, Spain, Hotel Porta Fira, January 17–18, 2017, pp. 159–171. SIAM (2017). https://doi.org/10.1137/1.9781611974768.13
7. de Boer, J., Derrida, B., Flyvbjerg, H., Jackson, A.D., Wettig, T.: Simple model of self-organized biological evolution. Phys. Rev. Lett. **73**, 906–909 (1994). https://doi.org/10.1103/PhysRevLett.73.906
8. Deger, A., Brandner, K., Flindt, C.: Lee-Yang zeros and large-deviation statistics of a molecular zipper. Phys. Rev. E **97**, 012115 (2018). https://doi.org/10.1103/PhysRevE.97.012115
9. Dickman, R.: Nonequilibrium lattice models: series analysis of steady states. J. Stat. Phys. **55**(5–6), 997–1026 (1989). https://doi.org/10.1007/BF01041076
10. Dickman, R., Garcia, G.J.M.: Absorbing-state phase transitions with extremal dynamics. Phys. Rev. E **71**, 066113 (2005). https://doi.org/10.1103/PhysRevE.71.066113
11. Harris, T.E.: Contact interactions on a lattice. Ann. Probab. **2**(6), 969–988 (1974). https://doi.org/10.1214/aop/1176996493
12. Hunter, D.L., Baker, G.A.: Methods of series analysis. I. Comparison of current methods used in the theory of critical phenomena. Phys. Rev. B **7**, 3346–3376 (1973). https://doi.org/10.1103/PhysRevB.7.3346
13. Hunter, D.L., Baker, G.A.: Methods of series analysis. III. Integral approximant methods. Phys. Rev. B **19**, 3808–3821 (1979). https://doi.org/10.1103/PhysRevB.19.3808
14. Inui, N.: Series expansion for a nonequilibrium lattice model with parity conservation. J. Phys. Soc. Jpn. **64**(7), 2266–2269 (1995). https://doi.org/10.1143/JPSJ.64.2266
15. Inui, N.: Distribution of poles in a series expansion of the asymmetric directed-bond percolation probability on the square lattice. J. Phys. A **31**(48), 9613–9620 (1998). https://doi.org/10.1088/0305-4470/31/48/001
16. Inui, N., Katori, M.: Catalan numbers in a series expansion of the directed percolation probability on a square lattice. J. Phys. A **29**(15), 4347–4364 (1996). https://doi.org/10.1088/0305-4470/29/15/010
17. Jensen, I., Dickman, R.: Time-dependent perturbation theory for nonequilibrium lattice models. J. Stat. Phys. **71**(1–2), 89–127 (1993). https://doi.org/10.1007/BF01048090
18. Jovanović, B., Buldyrev, S.V., Havlin, S., Stanley, H.E.: Punctuated equilibrium and "history-dependent" percolation. Phys. Rev. E 50, R2403–R2406 (1994). https://doi.org/10.1103/PhysRevE.50.R2403
19. Katori, M., Tsuchiya, T., Inui, N., Kakuno, H.: Baxter–Guttmann–Jensen conjecture for power series in directed percolation problem. Ann. Comb. **3**(2–4), 337–356 (1999). https://doi.org/10.1007/BF01608792. On combinatorics and statistical mechanics
20. Marsili, M.: Renormalization group approach to the self-organization of a simple model of biological evolution. EPL **28**(6), 385 (1994). https://doi.org/10.1209/0295-5075/28/6/002
21. Marsili, M., De Los Rios, P., Maslov, S.: Expansion around the mean-field solution of the Bak–Sneppen model. Phys. Rev. Lett. **80**, 1457–1460 (1998). https://doi.org/10.1103/PhysRevLett.80.1457
22. Meester, R., Znamenski, D.: Non-triviality of a discrete Bak–Sneppen evolution model. J. Stat. Phys. **109**(5–6), 987–1004 (2002). https://doi.org/10.1023/A:1020468325294. MR1937000
23. Peters, H., Regts, G.: On a conjecture of sokal concerning roots of the independence polynomial. Mich. Math. J. **68**(1), 33–35 (2018). https://doi.org/10.1307/mmj/1541667626
24. Pruessner, G.: Self-Organised Criticality: Theory, Models and Characterisation. Cambridge University Press, Cambridge (2012). https://doi.org/10.1017/CBO9780511977671
25. Tretyakov, A.Y., Inui, N., Konno, N.: Phase transition for the one-sided contact process. J. Phys. Soc. Jpn. **66**(12), 3764–3769 (1997). https://doi.org/10.1143/JPSJ.66.3764

26. Yang, C.N., Lee, T.D.: Statistical theory of equations of state and phase transitions. I. Theory of condensation. Phys. Rev. **87**, 404–409 (1952). https://doi.org/10.1103/PhysRev.87.404

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