

# Branching within branching: a general model for host-parasite co-evolution

Gerold Alsmeyer (joint work with Sören Gröttrup)

April 9, 2015

- 1 Model
- 2 The associated branching process in random environment (ABPRE)
- 3 Extinction results
- 4 Limit theorems in the case of survival

# Model description

## The basic ingredients

- a cell population (the hosts)
- a population of parasites colonizing the cells

## The basic assumptions

- cells form an ordinary Galton-Watson tree (GWT)
- parasites sitting in different cells multiply and share their progeny into daughter cells independently, **but**
- for parasites hosted by the same cell  $v$ , offspring numbers and sharing of progeny are **conditionally independent given the number of daughter cells of  $v$**

# Notational details: the cell population

- Ulam-Harris tree  $\mathbb{V} = \bigcup_{n \geq 0} \mathbb{N}^n$  with  $\mathbb{N}^0 = \{\emptyset\}$ .
- **cell population**: a GWT  $\mathbb{T} = \bigcup_{n \in \mathbb{N}_0} \mathbb{T}_n \subset \mathbb{V}$  with  $\mathbb{T}_0 = \{\emptyset\}$  and

$$\mathbb{T}_n := \{v_1 \dots v_n \in \mathbb{V} \mid v_1 \dots v_{n-1} \in \mathbb{T}_{n-1} \text{ and } 1 \leq v_n \leq N_{v_1 \dots v_{n-1}}\},$$

where  $N_v$  denotes the number of daughter cells of  $v$ .

- the  $N_v$ ,  $v \in \mathbb{V}$  are iid with common law  $(p_k)_{k \geq 0}$ , the **offspring distribution of cells**, having finite mean  $\nu = \sum_{k \geq 1} k p_k$ .
- the number of cells process:  $\mathcal{I}_n = \sum_{v \in \mathbb{T}_{n-1}} N_v$  (a GWP).

## Details: the parasites

- the number of parasites in cell  $\mathbf{v}$  are denoted by  $Z_{\mathbf{v}}$ .
- the number of parasites process is then defined by

$$\mathcal{Z}_n := \sum_{\mathbf{v} \in \mathbb{T}_n} Z_{\mathbf{v}}, \quad n \in \mathbb{N}_0.$$

- the set of contaminated cells:  $\mathbb{T}_n^* = \{\mathbf{v} \in \mathbb{T}_n : Z_{\mathbf{v}} > 0\}$ .
- the number of contaminated cells:  $\mathcal{I}_n^* = \#\mathbb{T}_n^*$ .
- cell counts with a specific number of parasites:

$$\mathcal{T}_n := (\mathcal{I}_{n,0}, \mathcal{I}_{n,1}, \mathcal{I}_{n,2}, \dots),$$

where  $\mathcal{I}_{n,k}$  gives the number of cells in generation  $n$  hosting  $k$  parasites.

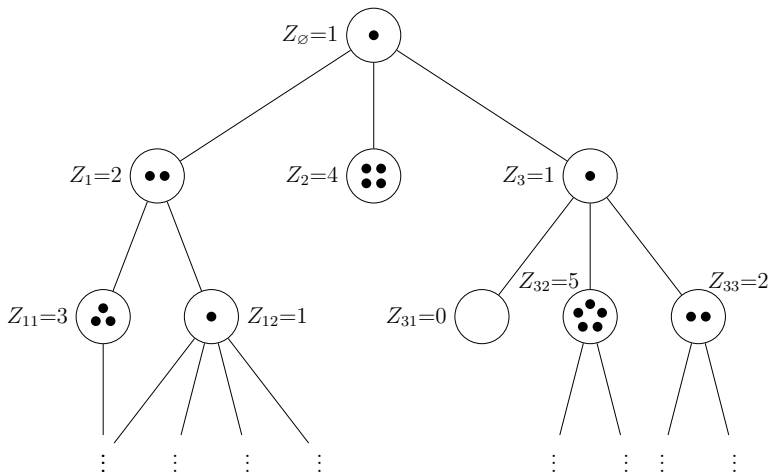


Figure : A typical realization of the first three generations of a BwBP.

# Reproduction of parasites

To describe **reproduction of parasites** consider those hosted by a cell  $v \in \mathbb{T}$  and suppose that

- $Z_v = z \geq 1$  and that
- $v$  has  $k$  daughter cells, labeled  $v_1, \dots, v_k$ , thus  $N_v = k$ .

For  $i = 1, \dots, z$ , let

$$X_{i,v}^{(\bullet,k)} := \left( X_{i,v}^{(1,k)}, \dots, X_{i,v}^{(k,k)} \right)$$

be iid copies of a random vector  $X^{(\bullet,k)} := (X^{(1,k)}, \dots, X^{(k,k)})$  with **arbitrary law on  $\mathbb{N}_0^k$  and independent of any other occurring  $v$ 's**. Then, given  $N_v = k$ , and for  $i = 1, \dots, Z_v = z$ ,

- $X_{i,v}^{(j,k)}$  equals the number of offspring of the  $i^{\text{th}}$  parasite in  $v$  that is shared into daughter cell  $v_j$ .

# Model parameters and basic assumptions

- $\mu_{j,k} = \mathbb{E}X^{(j,k)}$
- $\gamma = \mathbb{E}\mathcal{Z}_1 = \sum_{k \geq 1} \rho_k \sum_{j=1}^k \mu_{j,k}$ , the mean number of offspring per parasite, is supposed to be **positive and finite**, thus
- $\mu_{j,k} < \infty$  for all  $j \leq k$  and  $\mathbb{P}(N_v = 0) < 1$ .
- We further assume  $\rho_1 = \mathbb{P}(N_v = 1) < 1$ ,  $\mathbb{P}(\mathcal{Z}_1 = 1) < 1$ , and
- a positive chance for more than one parasite to be shared in the same daughter cell:

$$\rho_k \mathbb{P}(X^{(j,k)} \geq 2) > 0 \quad \text{for at least one } (j,k), 1 \leq j \leq k.$$



# A very short list of earlier contributions



M. Kimmel.

Quasistationarity in a branching model of division-within-division.

*In Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 157–164. Springer, New York, 1997.



V. Bansaye.

Proliferating parasites in dividing cells: Kimmel's branching model revisited.

*Ann. Appl. Probab.*, 18(3):967–996, 2008.



G. Alsmeyer and S. Gröttrup.

A host-parasite model for a two-type cell population.

*Adv. in Appl. Probab.*, 45(3):719–741, 2013.

# The notorious questions

- Extinction-explosion dichotomy for the number of parasites process  $(\mathcal{L}_n)_{n \geq 0}$ .
- Extinction-explosion dichotomy for the number of contaminated cells process  $(\mathcal{T}_n^*)_{n \geq 0}$
- $\{\mathcal{L}_n \rightarrow \infty\} = \{\mathcal{T}_n^* \rightarrow \infty\}$ ?
- Necessary and sufficient conditions for almost sure extinction of contaminated cells.
- Limit theorems for the relevant processes in the survival case (after normalization).

# The ABPRE

The associated branching process in random environment (ABPRE) is obtained by picking an infinite random cell-line (spine) in a size-biased version of  $\mathbb{T}$ :

# The construction (standard)

- Let  $(T_n, C_n)_{n \geq 0}$  be a sequence of iid random vectors independent of  $(N_v)_{v \in \mathbb{V}}$  and  $(X_{i,v}^{(\bullet, k)})_{k \geq 1, i \geq 1, v \in \mathbb{V}}$ .
- The law of  $T_n$  equals the size-biasing of the law of the  $N_v$ , i.e.

$$\mathbb{P}(T_n = k) = \frac{kp_k}{v}$$

for each  $n \in \mathbb{N}_0$  and  $k \in \mathbb{N}$ , and

$$\mathbb{P}(C_n = j | T_n = k) = \frac{1}{k}$$

for  $1 \leq j \leq k$ , which means that  $C_n$  has a uniform distribution on  $\{1, \dots, k\}$  given  $T_n = k$ .

# The ABPRE

The **random cell-line (spine)**  $(V_n)_{n \geq 0}$  is then recursively defined by  $V_0 = \emptyset$  and

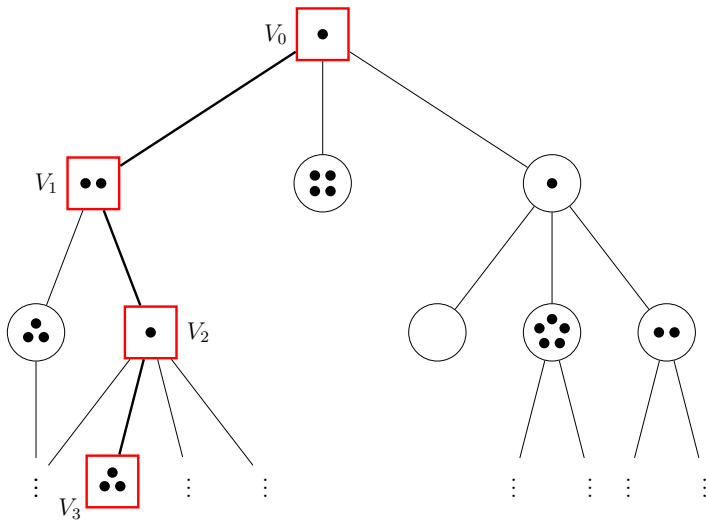
$$V_n := V_{n-1} C_{n-1}$$

for  $n \geq 1$ . Then

$$\emptyset =: V_0 \rightarrow V_1 \rightarrow V_2 \rightarrow \dots \rightarrow V_n \rightarrow \dots$$

provides us with a random cell line in  $\mathbb{V}$  (not picked uniformly) as depicted in the following picture.

# The ABPRE



# The number of parasites along the spine

Regarding the structure of **the number of parasites process along the spine**  $(Z_{V_n})_{n \geq 0}$ , the following lemma is fundamental.

## Lemma

Let  $(Z'_n)_{n \geq 0}$  be a BPPE with  $Z_\emptyset$  ancestors and iid environmental sequence  $\Lambda := (\Lambda_n)_{n \geq 0}$  taking values in  $\{\mathcal{L}(X^{(j,k)}) \mid 1 \leq j \leq k < \infty\}$  and such that

$$\mathbb{P}\left(\Lambda_0 = \mathcal{L}(X^{(j,k)})\right) = \frac{p_k}{v} = \frac{1}{k} \cdot \frac{kp_k}{v} = \mathbb{P}(C_0 = j, T_0 = k)$$

for all  $1 \leq j \leq k < \infty$ . Then  $(Z_{V_n})_{n \geq 0}$  and  $(Z'_n)_{n \geq 0}$  are equal in law.

# Definition of the ABPRE

We call the BPRE  $(Z'_n)_{n \geq 0}$  the **associated branching process in random environment (ABPRE)**. It is one of the major tools used in the study of the BwBP, and the following lemma provides a key relation between this process and its associated ABPRE.



# A key result

## Lemma

For all  $n, k, z \in \mathbb{N}_0$ ,

$$\mathbb{P}_z(Z'_n = k) = v^{-n} \mathbb{E}_z \mathcal{I}_{n,k}$$

and

$$\mathbb{P}_z(Z'_n > 0) = v^{-n} \mathbb{E}_z \mathcal{I}_n^*.$$

# Generating functions

For  $n \in \mathbb{N}$  and  $s \in [0, 1]$

$$\mathbb{E}(s^{Z'_n} | \Lambda) = g_{\Lambda_0} \circ \dots \circ g_{\Lambda_{n-1}}(s)$$

is the **quenched generating function** of  $Z'_n$  with iid  $g_{\Lambda_n}$  and  $g_\lambda$  defined by

$$g_\lambda(s) := \mathbb{E}(s^{Z'_1} | \Lambda_0 = \lambda) = \sum_{n \geq 0} \lambda_n s^n$$

for any distribution  $\lambda = (\lambda_n)_{n \geq 0}$  on  $\mathbb{N}_0$ . Moreover,

$$\mathbb{E}g'_{\Lambda_0}(1) = \mathbb{E}Z'_1 = \sum_{1 \leq j \leq k} \frac{\rho_k}{v} \mathbb{E}X^{(j,k)} = \frac{\mathbb{E}\mathcal{Z}_1}{v} = \frac{\gamma}{v} < \infty, \quad (1)$$

where  $\gamma = \mathbb{E}\mathcal{Z}_1$ .

# Regimes of the ABPRE

It is well-known that  $(Z'_n)_{n \geq 0}$  **survives with positive probability** iff

$$\mathbb{E} \log g'_{\Lambda_0}(1) > 0 \quad \text{and} \quad \mathbb{E} \log^-(1 - g_{\Lambda_0}(0)) < \infty.$$

Recall that  $\gamma < \infty$  is assumed and that there exists  $1 \leq j \leq k < \infty$  such that  $p_k > 0$  and  $\mathbb{P}(X^{(j,k)} \neq 1) > 0$ , which ensures that  $\Lambda_0 \neq \delta_1$  with positive probability. The ABPRE is called **supercritical**, **critical** or **subcritical** if  $\mathbb{E} \log g'_{\Lambda_0}(1) > 0$ ,  $= 0$  or  $< 0$ , respectively.

# Regimes of the ABPRE

The subcritical case further divides into the three subregimes when  $\mathbb{E}g'_{\lambda_0}(1) \log g'_{\lambda_0}(1) < 0, = 0,$  or  $> 0$ , respectively, called **strongly, intermediately** and **weakly subcritical case**. The quite different behavior of the process in each of the three subregimes is shown by the limit results derived in



J. Geiger, G. Kersting, and V. A. Vatutin.

Limit theorems for subcritical branching processes in random environment.

*Ann. Inst. H. Poincaré Probab. Statist.*, 39(4):593–620, 2003.

# Notation

For  $\vec{s} = (s_0, s_1, \dots) \in \mathbf{N} = \{(x_i)_{i \geq 0} \in \mathbb{N}_0^\infty \mid x_i > 0 \text{ finitely often}\}$  and  $z \in \mathbb{N}_0$ , we use

- $\mathbb{P}_{\vec{s}}$  for probabilities conditioned upon the event that the initial generation consists of  $s_k$  cells hosting exactly  $k$  parasites for  $k = 0, 1, \dots$ , i.e.  $\mathcal{T}_{0,k} = s_k$  for  $k = 0, 1, \dots$
- $\mathbb{P}_z$  for probabilities given that initially there is one cell which contains  $z$  parasites, i.e.  $N_\emptyset = 1$  and  $Z_\emptyset = z$ .
- $\mathbb{P} = \mathbb{P}_1$ .

# The number of parasites process

## Theorem (Extinction-explosion principle)

*The parasite population of a BwBP either dies out or explodes, i.e. for all  $\vec{s} \in \mathbf{N} = \{(x_i)_{i \geq 1} \in \mathbb{N}_0^\infty \mid x_i > 0 \text{ finitely often}\}$*

$$\mathbb{P}_{\vec{s}}(\mathcal{Z}_n \rightarrow 0) + \mathbb{P}_{\vec{s}}(\mathcal{Z}_n \rightarrow \infty) = 1.$$

# The number of contaminated cells process

## Theorem

Let  $\mathbb{P}(\text{Surv}) > 0$ .

- (a) If  $\mathbb{P}_2(\mathcal{I}_1^* \geq 2) > 0$ , then  $\mathbb{P}_z^*(\mathcal{I}_n^* \rightarrow \infty) = 1$  and thus  
 $\text{Ext} = \{\sup_{n \geq 0} \mathcal{I}_n^* < \infty\}$   $\mathbb{P}_z$ -a.s. for all  $z \in \mathbb{N}$ .
- (b) If  $\mathbb{P}_2(\mathcal{I}_1^* \geq 2) = 0$ , then  $\mathbb{P}_z^*(\mathcal{I}_n^* = 1 \text{ f.a. } n \geq 0) = 1$  for all  
 $z \in \mathbb{N}$ .

# Almost sure extinction of contaminated cells

## Theorem

(a) If  $\mathbb{P}_2(\mathcal{I}_1^* \geq 2) = 0$ , then  $\mathbb{P}(\text{Ext}) = 1$  if, and only if,

$$\mathbb{E} \log \mathbb{E}(\mathcal{Z}_1 | N_\emptyset) \leq 0 \quad \text{or} \quad \mathbb{E} \log^- \mathbb{P}(\mathcal{Z}_1 > 0 | N_\emptyset) = \infty.$$

(b) If  $\mathbb{P}_2(\mathcal{I}_1^* \geq 2) > 0$ , then the following statements are equivalent:

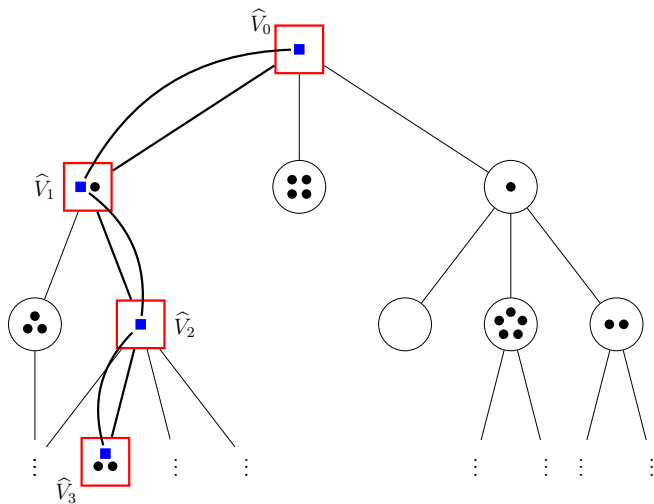
- (i)  $\mathbb{P}(\text{Ext}) = 1$ .
- (ii)  $\mathbb{E} \mathcal{I}_n^* \leq 1$  for all  $n \in \mathbb{N}_0$ .
- (iii)  $\sup_{n \in \mathbb{N}_0} \mathbb{E} \mathcal{I}_n^* < \infty$ .
- (iv)  $v \leq 1$ , or

$$v > 1, \quad \mathbb{E} \log g'_{\Lambda_0}(1) < 0 \quad \text{and} \quad \inf_{0 \leq \theta \leq 1} \mathbb{E} g'_{\Lambda_0}(1)^\theta \leq \frac{1}{v}.$$



Now consider the survival case.

## A second size-biasing



## A KS-type result for the number of contaminated cells

## Theorem

$(v^{-n} \mathcal{I}_n^*)_{n \geq 0}$  is a nonnegative supermartingale with respect to the filtration  $(\mathcal{F}_n)_{n \geq 0}$  and therefore almost surely convergent to an integrable random variable  $L$  as  $n \rightarrow \infty$ . Furthermore,

- (a)  $\mathbb{P}(L = 0) < 1$  iff the following three conditions hold true:
- (i)  $v > 1$ .
  - (ii)  $\mathbb{E}N \log N < \infty$ .
  - (iii)  $\mathbb{E} \log g'_{\Lambda_0}(1) > 0$  and  $\mathbb{E} \log^-(1 - g_{\Lambda_0}(0)) < \infty$ .
- (b)  $\mathbb{P}(L = 0) < 1$  implies  $\{L = 0\} = \text{Ext}$  a.s.

# Heyde-Seneta norming

## Theorem

*If  $\mathbb{E} \log g'_{\Lambda_0}(1) > 0$  and  $\mathbb{E} \log^-(1 - g_{\Lambda_0}(0)) < \infty$ , then there exists a sequence  $(c_n)_{n \geq 0}$  in  $(0, \infty)$  such that  $c_{n+1}/c_n \rightarrow v$  and  $c_n^{-1} \mathcal{T}_n^*$  converges a.s. to a finite random variable  $L^*$  satisfying  $\mathbb{P}(L^* > 0) = \mathbb{P}(\text{Surv})$ . Furthermore, the sequence  $(c_n)_{n \geq 0}$  is a proper Heyde-Seneta norming for  $(\mathcal{T}_n)_{n \geq 0}$  as well.*

# A KS-type result for the number of parasites

The normalized number of parasites process

$$W_n = \gamma^{-n} \mathcal{Z}_n, \quad n \geq 0$$

is a martingale.

## Theorem

*The expectation of  $W$  is either 0 or 1, and*

$$\mathbb{E}W = 1 \quad \text{iff} \quad \mathbb{E}\mathcal{Z}_1 \log \mathcal{Z}_1 < \infty \quad \text{and} \quad \mathbb{E} \left( \frac{g'_{\Lambda_0}(1)}{\gamma} \log \frac{g'_{\Lambda_0}(1)}{\gamma} \right) < 0.$$

*in which case  $\mathbb{P}(W > 0) = \mathbb{P}(\text{Surv})$ .*

# Want to know more?



[G. Alsmeyer and S. Gröttrup.](#)

Branching within branching I: Extinction properties, 2015.  
Preprint (19 pages).



[G. Alsmeyer and S. Gröttrup.](#)

Branching within branching II: Limit theorems, 2015.  
Preprint (38 pages).