Combinatorics, Probability and Computing (2020), **29**, pp. 455–484
doi:10.1017/S0963548319000403

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Minimax functions on Galton–Watson trees

James B. Martin and Roman Stasiński

Department of Statistics, University of Oxford, Oxford OX1 3LB, UK
Email: martin@stats.ox.ac.uk

(Received 22 July 2018; revised 20 August 2019; first published online 6 December 2019)

Abstract

We consider the behaviour of minimax recursions defined on random trees. Such recursions give the value of a general class of two-player combinatorial games. We examine in particular the case where the tree is given by a Galton–Watson branching process, truncated at some depth $2n$, and the terminal values of the level $2n$ nodes are drawn independently from some common distribution. The case of a regular tree was previously considered by Pearl, who showed that as $n \to \infty$ the value of the game converges to a constant, and by Ali Khan, Devroye and Neininger, who obtained a distributional limit under a suitable rescaling.

For a general offspring distribution, there is a surprisingly rich variety of behaviour: the (unrescaled) value of the game may converge to a constant, or to a discrete limit with several atoms, or to a continuous distribution. We also give distributional limits under suitable rescalings in various cases.

We also address questions of *endogeny*. Suppose the game is played on a tree with many levels, so that the terminal values are far from the root. To be confident of playing a good first move, do we need to see the whole tree and its terminal values, or can we play close to optimally by inspecting just the first few levels of the tree? The answers again depend in an interesting way on the offspring distribution.

We also mention several open questions.

2010 MSC Codes: Primary 60J80; Secondary 60G99, 91A46

1. Introduction

In this paper we consider the behaviour of minimax recursions defined on random trees. Consider a finite rooted tree with depth m. We will call the root 'level 0', the children of the root 'level 1', and so on. Suppose every node at levels $0, 1, \ldots, m-1$ has at least one child; the nodes at level m are all leaves. Suppose every leaf node (*i.e.* every node at level m) has some real value associated to it. Then recursively propagate the values towards the root in a minimax way: each node at an odd level gets a value which is the max of the values of its children, and each node at an even level gets a value which is the min of the values of its children.

This minimax procedure has a natural interpretation in terms of a two-player game. Two players alternate turns; a token starts at the root, and a move of the game consists of moving the token from its current node to one of the children of that node. The leaf nodes are terminal positions; the outcome of the game is the value associated to the leaf node where the game ends. Player 1 is trying to minimize this outcome, and player 2 is trying to maximize it. The outcome of the game with 'optimal play' is the value associated to the root.

Suppose the terminal values are random, drawn independently from some common distribution.

© Cambridge University Press 2019

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Pearl [10] considered the case where the tree is regular (every non-leaf node has d children for some $d \geqslant 2$) and the terminal values are independent and uniformly distributed on the interval $[0, 1]$. For simplicity assume that the depth of the tree is even; write W_{2n} for a random variable representing the value at the root of a tree of depth $2n$. Pearl showed that W_{2n} converges in distribution to a constant as $n \to \infty$. This result was refined by Ali Khan, Devroye and Neininger [2], who derived an asymptotic distribution for W_{2n} after appropriate rescaling.

In this paper we consider the case where the tree is given by a Galton–Watson branching process, truncated at level $2n$. This generalization leads to a surprisingly rich variety of behaviour, depending on the offspring distribution of the branching process. For example, the limiting distribution of W_{2n} may be concentrated at a single point (as in the regular case), or may now have several atoms, or may even be continuous.

There is also a rich interplay between the two sources of randomness now present in the model (the tree itself, and the terminal values at the leaves). Suppose we play the game on a tree with many levels, so that the terminal values are far from the root. In order to be confident of playing a good first move, do we need to see the whole tree and terminal values, or can we play close to optimally by inspecting just the structure of the first few levels of the tree? Such questions can be formulated precisely in terms of the *endogeny* property for certain recursive tree processes, as introduced by [1]. The answers again depend in an interesting way on the offspring distribution.

Such questions concerning the relative importance of local tree structure and terminal values are of considerable interest in understanding the effectiveness of certain tree-search algorithms such as *Monte Carlo tree search* (MCTS); see [6] for a survey. MCTS has famously been applied in recent years to games such as go, where it provided a considerable increase in playing strength [7] even before being allied with powerful deep learning techniques [12]. For some games, simple versions of these algorithms, without local evaluation functions, and with only very crude input from the terminal values (given for example by 'random rollouts' through unexplored parts of the tree), are nonetheless able to converge quickly towards good lines of play. Understanding which aspects of a game's structure make such convergence possible is an interesting challenge both in theoretical and in practical terms.

Our main results concerning distributional limits are presented in the next section. In Section 2 we discuss a range of examples and mention some open problems. The results about endogeny are given in Section 3. The main proofs are given in Sections 4 and 5.

Before that we mention some recent related work. Broutin, Devroye and Fraiman [4] consider recursive distributional equations (including those of minimax type) defined on Galton–Watson trees conditioned to have a given total size n. Holroyd and Martin [8] consider minimax-type games (and various *misère* and asymmetric variants) defined on (perhaps infinite) Galton–Watson trees, with particular emphasis on the nature of phase transitions for the outcomes of the game as the underlying offspring distribution varies (see Section 2 for further comments). Note that in both [4] and [8], unlike in the case of this paper, the offspring distribution puts positive weight at 0, so that there are leaves close to the root.

Similar questions arise in the context of random AND/OR trees and random Boolean functions. For example the model of Pemantle and Ward [11] involves a regular tree in which each node independently is a max or a min with equal probability; see Section 2.2 for comments on the relation to a particular case of our model. See for example Broutin and Mailler [5] for a variety of recent results in a more general setting, and many relevant references.

1.1 Main results

Consider a Galton–Watson tree with an offspring distribution with mass function $p_1, p_2, p_3, \ldots$ on $\{1, 2, 3, \ldots\}$ (note that every individual has at least one child). Let

$$G(x) = \sum_{k=0}^{\infty} p_k x^k$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

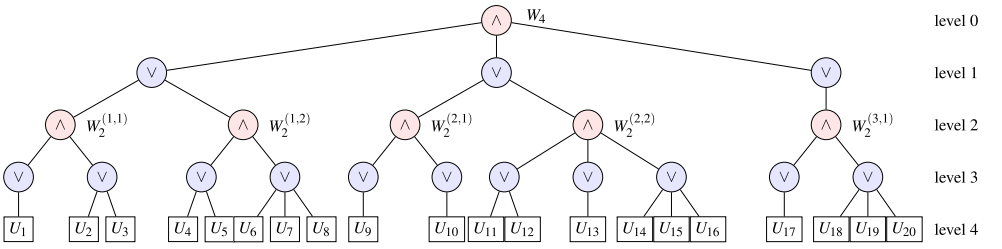


Figure 1. An example of a minimax tree, with 4 levels. Here all non-leaf nodes have 1, 2 or 3 children.

be the probability generating function of the offspring distribution (which is a strictly increasing function mapping [0, 1] to [0, 1] bijectively). We will also write throughout

$$R(x) = 1 - G(x)$$

and

$$f(x) = R(R(x)). \tag{1.1}$$

Truncate the tree at level $2n$, so that all the vertices at level $2n$ are leaves. Let the terminal values associated to the leaves be i.i.d. uniform on $[0, 1]$ (independently of the structure of the tree). Recursively, assign values to the internal nodes of the tree (in particular to the root) using the minimax procedure defined above. See Figure 1 for an illustration.

(Note that there is nothing particularly special about uniform boundary conditions. By a simple rescaling we can map between this case and the case of i.i.d. boundary values from any other continuous distribution. Later we will also consider discrete boundary values, for example those taking values only 0 and 1, where we can interpret 0 as a win for the first player, and 1 as a win for the second player.)

We let W_{2n} denote the random variable associated with the root of a tree of depth $2n$. Then we have a distributional recursion,

$$W_{2n} \stackrel{d}{=} \min_{1 \leqslant i \leqslant M} \max_{1 \leqslant j \leqslant M_i} W_{2n-2}^{(i,j)}, \tag{1.2}$$

where M and $M_1, M_2, M_3, \ldots$ are i.i.d. draws from the offspring distribution, and $W_{2n-2}^{(i,j)}$, $i, j \in \mathbb{N}$, are i.i.d. copies of the random variable W_{2n-2}, independent of M and $\{M_i\}$.

A simple generating function computation (see the beginning of Section 4) gives

$$\mathbb{P}(W_{2n} \leqslant x) = f(\mathbb{P}(W_{2n-2} \leqslant x)), \tag{1.3}$$

where f is defined at (1.1). So to look at the behaviour of W_{2n} for large n, we will be interested in the function f and in particular its fixed points.

We begin with the results for the case of a regular tree.

Theorem 1.1. *Suppose $p_d = 1$ for some $d \geqslant 2$.*

(a) (Pearl [10])

$$W_{2n} \stackrel{d}{\to} q \ \text{ as } n \to \infty,$$

where q is the unique fixed point in $(0, 1)$ of the function $f_{d\text{-reg}}$ defined by

$$f_{d\text{-reg}}(x) = 1 - (1 - x^d)^d.$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

(b) (Ali Khan, Devroye and Neininger [2]) *Let* $\xi = f'_{d\text{-reg}}(q)$. *Then*

$$\xi^n(W_{2n} - q) \xrightarrow{d} W \quad \text{as } n \to \infty,$$

where W *has a continuous distribution function* F_W *which satisfies* $F_W(x) = f_{d\text{-reg}}(F_W(x/\xi))$.

Now we will consider general offspring distributions. Since G is increasing and bijective as a function from $[0, 1]$ to $[0, 1]$, we have that $R = 1 - G$ is decreasing and bijective, and $f = R \circ R$ is again increasing and bijective. Also G is analytic on $[0, 1)$, so that f is analytic on $(0, 1)$.

We will be particularly interested in fixed points of the function f. The function R itself has a single fixed point, which is obviously also a fixed point of f. Otherwise the fixed points of f come in pairs: if q is one then so is $R(q)$. One such pair are the points 0 and 1. We will say that a fixed point q of f is *unstable from the right* if $q < 1$ and $\lim_{\varepsilon\to 0} \lim_{n\to\infty} f^n(q + \varepsilon) > q$; similarly it is *unstable from the left* if $q > 0$ and $\lim_{\varepsilon\to 0} \lim_{n\to\infty} f^n(q - \varepsilon) < q$.

For a regular tree, Theorem 1.1 tells us that the distribution of W_{2n} converges to a constant. For general distributions, we still have convergence in distribution, but now we may have a 'genuinely random outcome' in the limit as the tree becomes large; the limiting distribution may have more than one atom (and in some surprising cases, the distribution of W_{2n} can simply be the same uniform distribution for all n).

Theorem 1.2. $W_{2n} \xrightarrow{d} W$ *as* $n \to \infty$, *for some random variable* W. *There are two cases.*

(a) *If* f *is the identity function, then* $W_{2n} \sim U[0, 1]$ *for all* n.
(b) *Otherwise, let* Q *be the set of fixed points of* f *which are unstable from at least one side. Then* W *is discrete and has atoms precisely at the elements of* Q.
For $q \in Q$, *define*

$$q_- = \begin{cases} \sup\{x : x < q, x = f(x)\} & \text{if } q > 0 \text{ and } q \text{ is unstable from the left,} \\ q & \text{otherwise,} \end{cases}$$
$$q_+ = \begin{cases} \inf\{x : x > q, x = f(x)\} & \text{if } q < 1 \text{ and } q \text{ is unstable from the right,} \\ q & \text{otherwise.} \end{cases} \tag{1.4}$$

Then $\mathbb{P}(W = q) = q_+ - q_-$.

It is not hard to show that $x \in Q$ if and only if $R(x) \in Q$. Hence again the atoms of the distributional limit W come in pairs, with the possible exception of the fixed point of R. In Section 2.1, we comment in particular on the case where W has atoms at 0 and 1.

For $q \in (0, 1)$, we may write (1.4) alternatively as

$$q_- = \lim_{\varepsilon\to 0} \lim_{n\to\infty} f^n(q - \varepsilon) \quad \text{and} \quad q_+ = \lim_{\varepsilon\to 0} \lim_{n\to\infty} f^n(q + \varepsilon)$$

(this follows straightforwardly from the monotonicity and continuity of f).

In the next results we consider fluctuations around the atoms of the limiting distributions obtained in Theorem 1.2(b). The appropriate rescaling around a point $q \in Q$ depends on the derivative of $\xi = f'(q)$. If $q \in Q$ then we must have $\xi \geqslant 1$.

Theorem 1.3. *Consider the model defined by* (1.2). *Assume that* f *is not the identity function and let* Q *be the set of fixed points of* f *unstable from at least one side.*
Let $q \in Q$. *Define* q_- *and* q_+ *as at* (1.4), *and let* $\xi = f'(q)$.

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

(a) *If* $1 < \xi < \infty$, *then*

$$\mathcal{L}(\xi^n(W_{2n} - q) \mid W_{2n} \in [q_-, q_+]) \to \mathcal{L}(V) \text{ as } n \to \infty,$$

where V *is a random variable with a continuous distribution function.*

(b) *Suppose* $\xi = 1$, *and* $k \geqslant 2$ *is such that* $f^{(r)}(q) = 0$ *for* $1 < r < k$ *and* $f^{(k)}(q) \neq 0$. *Then*

$$\mathcal{L}(n^{1/(k-1)}(W_{2n} - q) \mid W_{2n} \in [q_-, q_+]) \to \mathcal{L}(V),$$

where for

$$a = \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{1/(k-1)}$$

we have

$$V = \begin{cases} a & \text{with probability } \dfrac{q_+ - q}{q_+ - q_-}, \\ -a & \text{with probability } \dfrac{q - q_-}{q_+ - q_-}. \end{cases}$$

(c) *If* $\xi = \infty$, *then* $q \in \{0, 1\}$. *Assume now that*

$$\mathbb{E}(M\mathbb{I}_{M \leqslant n}) = \sum_{k=1}^{n} kp_k \sim cn^{\rho} \text{ as } n \to \infty \tag{1.5}$$

for some $c > 0$ *and* $\rho \in (0, 1)$, *where* M *is distributed according to the offspring distribution of the Galton–Watson tree, and let* $K = \min\{i : p_i \neq 0\}$. *Then* $K < 1/(1-\rho)$, *and* $|f(t) - q| \sim C|t - q|^{K(1-\rho)}$ *as* $t \to q$ *for some* $C > 0$. *Moreover,*

$$\mathcal{L}(-[K(1-\rho)]^n \log|W_{2n} - q| \mid W_{2n} \in [q_-, q_+]) \to \mathcal{L}(Y),$$

where Y *is a random variable such that* $\mathbb{P}(Y \in (0, \infty)) = 1$.

The scaling limits in part (a) are the closest ones to the result for the regular tree from Theorem 1.1. Note that when q is an endpoint of the interval, the limiting distribution V is now one-sided, supported on $(0, \infty)$ when $q = 0$ and on $(-\infty, 0)$ when $q = 1$.

For part (b), recall that f is analytic on $(0, 1)$ so certainly if $q \in (0, 1)$, such a k exists. Conceivably, there might be no such k in some cases where $q = 0$ or $q = 1$ (although we know of no example where analyticity fails at 0 or 1 except when the derivative is infinite).

On the other hand, many cases with $\xi = \infty$ are not covered by part (c). It seems challenging to describe all possible asymptotics; however, the assumption (1.5) is satisfied for an important class of power-law distributions with infinite mean, satisfying $\mathbb{P}(X > x) \sim x^{1-\alpha}$ with $\alpha \in (1, 2)$.

2. Examples, discussion and open questions

Our final main results, concerning the endogeny property, will be stated in Section 3. Before that, we discuss a variety of examples illustrating the results of Theorems 1.2 and 1.3.

First consider a case where each node has 1 or 3 children. This simple family already displays an interesting range of behaviours. Let $p_1 = p$ and $p_3 = 1 - p$, for $p \in [0, 1]$. In Figure 2, we plot the function $f(x) - x$ for $x \in [0, 1]$, for a variety of values of p. Fixed points of f correspond to zeros of the curve. A crossing from negative to positive corresponds to an unstable fixed point.

When $p < 0.5$, the points 0 and 1 are stable and there is a unique unstable fixed point in $(0, 1)$, just as in the case of a regular tree; W_{2n} converges to a constant. At $p = 0.5$, we have

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

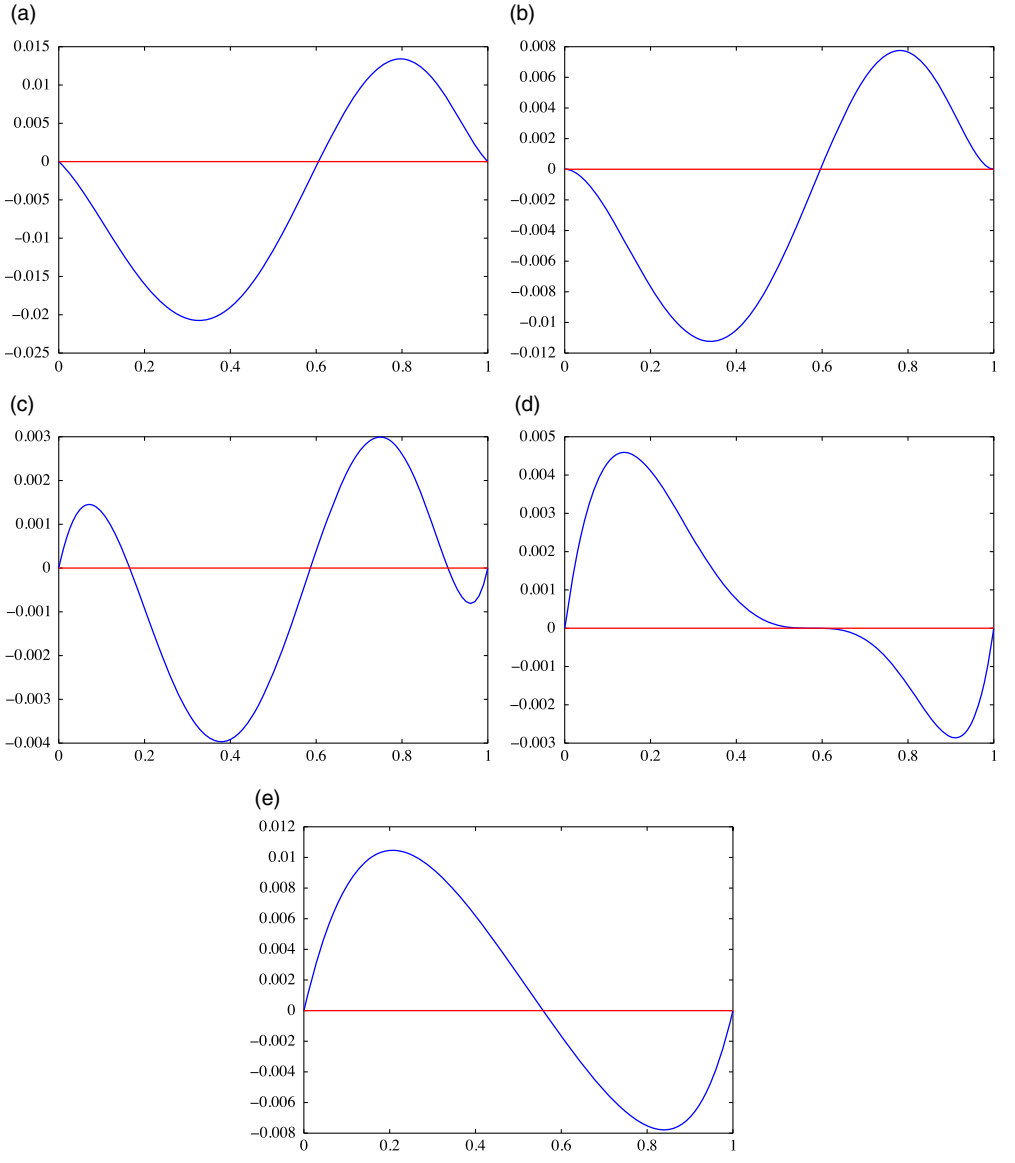


Figure 2. The function $f(x) - x$ for $x \in [0, 1]$ for the family of distributions with $p_1 = p$ and $p_3 = 1 - p$, with (a) $p = 0.45$, (b) $p = 0.5$, (c) $p = 0.55$, (d) $p = 0.598$, (e) $p = 0.7$.

$f'(0) = f'(1) = 1$; the slope of $f'(x) - x$ at 0 and 1 is 0, but the points are still stable. For $p > 0.5$, the points 0 and 1 are unstable, and the limiting distribution of W in Theorem 1.2 puts positive mass at 0 and 1. At first, there is also positive mass at another fixed point in (0, 1). However, above a critical point at roughly $p = 0.598$, two of the fixed points disappear, leaving only a stable fixed point in (0, 1), and the limiting distribution is concentrated only on the points 0 and 1.

Some further illustrative examples are shown in Figure 3.

For these distributions (the second and third are only approximate), we see points $q \in Q$ with $f'(q) = 1$, and so the rescalings of Theorem 1.3(b), which are polynomial rather than exponential, apply. In Figure 3(a) the relevant fixed points are at 0 and 1, and in Figure 3(b) they appear as 'touchpoints' in the graph of $f(x) - x$ so they are unstable from one side only; in these cases only

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

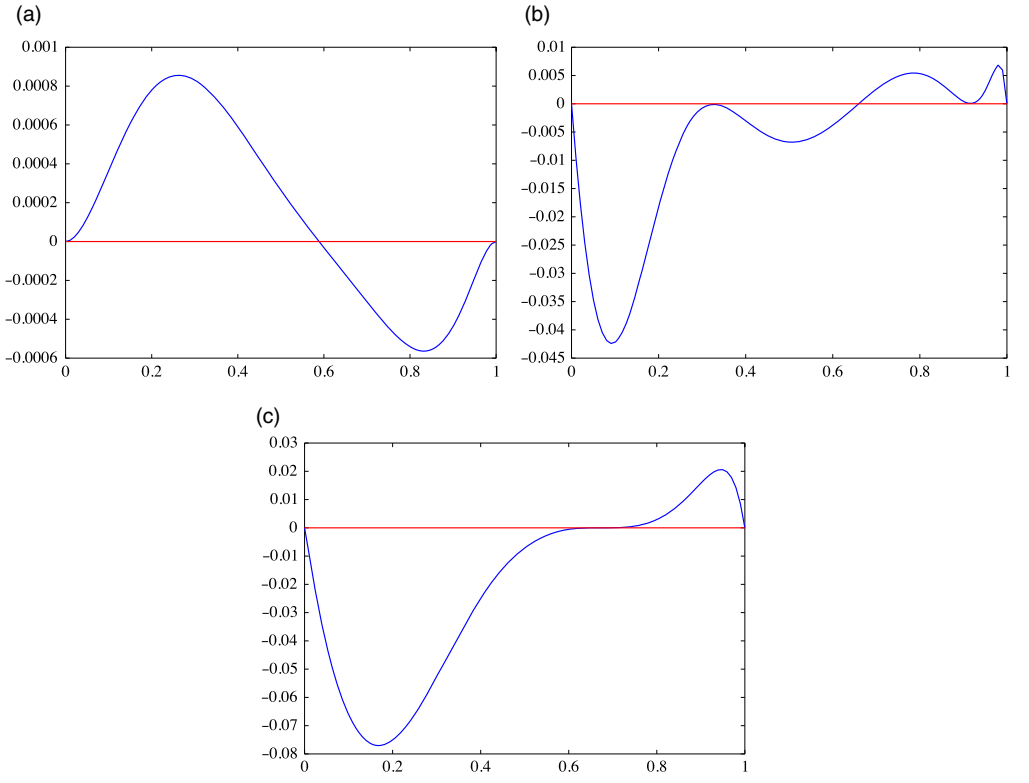


Figure 3. The function $f(x) - x$ for $x \in [0, 1]$ in three further cases: (a) $p_1 = 0.5$, $p_2 = 0.25$, $p_4 = 0.25$, (b) $p_2 = 0.783$, $p_{30} = 0.217$, (c) $p_2 = 0.705$, $p_{12} = 0.295$.

one of the points a and $-a$ in Theorem 1.3(b) receives positive mass. By contrast, in Figure 3(c), the point of inflection gives a fixed point which is unstable on both sides.

In passing, we mention briefly another interesting aspect of some of the above examples, concerning phase transitions. As we vary the offspring distribution, we see points where, for example, the number of atoms of the limiting distribution W changes. Often, the transition can be *continuous*: as the offspring distribution is varied, an existing atom may split into several new atoms (as may happen when a point of inflection occurs as in Figure 2(d) or Figure 3(c)), or new atoms may appear whose weight grows continuously from 0 (such as happens at the points 0 and 1 in Figure 2(b)). On the other hand, we can also see *discontinuous* transitions in cases such as Figure 3(b): one can perturb the offspring distribution in an arbitrarily small way to remove the 'touchpoints' seen there, so that the atoms of W inside $(0, 1)$ disappear and all their mass jumps to the endpoints 0 and 1. Such ideas, expressed only vaguely here, are studied in a closely related context in [8].

2.1 Atoms at endpoints

The limiting distribution W in Theorem 1.2 may have atoms at 0 and 1. We note the following simple criterion.

Proposition 2.1. *Let μ be the mean of the offspring distribution.*

(i) *If $p_1\mu < 1$ then $\mathbb{P}(W = 0) = \mathbb{P}(W = 1) = 0$.*
(ii) *If $p_1\mu > 1$ (including the case $p_1 > 0$ and $\mu = \infty$) then $\mathbb{P}(W = 0) > 0$ and $\mathbb{P}(W = 1) > 0$.*

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

If $p_1\mu = 1$, or if $p_1 = 0$ and $\mu = \infty$, either case is possible. The proof of the result is straightforward. Since $f = R \circ R$ we have $f'(x) = R'(R(x))R'(x)$. Then since $R(0) = 1$ and $R(1) = 0$, and since $R' = -G'$, we have $f'(0) = f'(1) = G'(0)G'(1) = \mu p_1$ (assuming $\mu < \infty$); and we know that a fixed point q of f is an atom of W if $f'(q) > 1$, and not if $f'(q) < 1$.

There is a rather direct interpretation of the condition $p_1\mu > 1$ in terms of the Galton–Watson tree and the play of the game. Consider the set of paths in the tree, starting at the root, with the following property: every vertex along the path at an odd level has only one child. The union of these paths gives a subtree containing the root. For a vertex at an even level (such as the root), the expected number of grandchildren in the subtree is $p_1\mu$, since the vertex itself has an average of μ children, and each of those has precisely one child with probability p_1. Considering only even levels, this then gives a branching process with mean offspring $p_1\mu$; if $p_1\mu > 1$, then this branching process is supercritical and survives for ever with positive probability. In that case, by keeping the game within this tree, the first player can ensure that the second player never has any choice at all; all the second player's moves are forced. For the game truncated at level $2n$, the first player can choose between all the nodes at level $2n$ which are within the subtree; from this it can be shown that $\mathbb{P}(W = 0)$ is at least as big as the probability that this branching process survives.

2.2 The case $f(x) \equiv x$, and related open questions

Suppose the offspring distribution is such that f is the identity function. Then from (1.3), if we put independent values at the leaves from any given distribution, then the value at the root has that same distribution (hence the statement in Theorem 1.2(a)). Perhaps surprisingly, this property is not restricted to the trivial case where $p_1 = 1$.

Here are some families of examples where $f = R \circ R = (1 - G) \circ (1 - G)$ is the identity (*i.e.* R is an *involution*).

(a) Any geometric distribution: if $p_k = p(1-p)^{k-1}$ for $p \in (0, 1)$, then $G(x) = p/(1 - (1 - p)x)$ and so $R(x) = 1 - x/(1 - (1 - p)x)$, and one can check $f(x) = x$.
(b) Let $G(x) = [1 - (1 - x)^{1/n}]^n$, for $n = 1, 2, 3, \ldots$. Via a binomial expansion, one can express G as a power series expansion with non-negative coefficients, and $G(1) = 1$, so G is indeed a probability generating function. The coefficient of x^k is non-zero for $k \geqslant n$.
(c) Let $G(x) = 1 - (1 - x^n)^{1/n}$, for $n = 1, 2, 3, \ldots$. Again G has a power series expansion with non-negative coefficients summing to 1. The coefficient of x^k is non-zero when k is a multiple of n.

These are far from the only cases. For a general source of examples, consider function $S(x, y)$ from $[0, 1]^2$ to $[0, 1]$ which is symmetric, increasing in each coordinate, and has $S(1, 0) = S(0, 1) = 0$. If we define a function R by setting $S(x, y) = 0$ and writing $y = R(x)$, then R is indeed an involution. Some such functions R have power series expansions, and in some of those cases $G = 1 - R$ has all coefficients positive, as needed for a probability generating function. For example, $S(x, y) = y^2 + y + x^2 + x - 2 = 0$ gives $R(x) = [\sqrt{9 - 4x - 4x^2} - 1]/2$, in which case one can obtain straightforwardly that $G = 1 - R$ is a generating function.

We note several questions that it might be interesting to understand further.

(1) Can one describe in some nice way the class of all distributions for which f is the identity? For the class of examples described in the previous paragraph, can one describe nicely which functions $S(x, y)$ lead to functions R which have power series expansions, and then which of those yield a generating function G?
(2) Are the geometric distributions in example (a) above the only such distributions with finite mean? More generally, what types of tail decay are possible? For (a), the tail $\sum_{r=k}^{\infty} p_r$ of course decays exponentially in k, while for (b) and (c) it decays as $k^{-1/n}$.

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

(3) Are there direct probabilistic arguments explaining the fact that f becomes the identity in these cases, in terms of the underlying process on the tree?

One case where it is possible to make such an argument is the $n = 2$ case in (b) above. Here p_k is the probability that the cluster containing the origin has size k for critical percolation on the binary tree (these coefficients are closely related to the Catalan numbers). Having made this identification, one can connect the minimax recursion on our random tree to an analogous recursion in the model studied by Pemantle and Ward [11], of a binary tree in which each node independently is a max or a min with probability 1/2 each.

We end this section with two further open questions about the form of f in more general cases.

(4) Can f have an arbitrarily large number of fixed points in $[0, 1]$?
(5) Can f have infinitely many fixed points in $[0, 1]$ (without being equal to the identity)? Since f is analytic on $(0, 1)$, this would require the set of fixed points to accumulate at 0 and at 1.

3. Endogeny

Suppose we play the game on a tree where the depth $2n$ is large, so that the boundary values are far from the root. To be confident of making a good first move, do we need to see a large part of the structure of the tree, and the boundary values? Or can we play close to optimally by inspecting just the structure of the first few levels of the tree? This is a so-called *endogeny* question [1]. The answer to this question again depends on the offspring distribution and the distribution of the boundary values.

To formalize the question, first define an operator on distributions corresponding to the recursion given by (1.2). For a distribution μ on $[0, 1]$, let $T(\mu)$ be the distribution of the left-hand side of (1.2) when the random variables $W_{2n-n}^{(i,j)}$ on the right-hand side are i.i.d. with distribution μ. Equivalently, rewriting (1.3), $T(\mu)[0, y] = f(\mu[0, y])$ for all y.

We will be interested in fixed points of T. For example, for offspring distributions such that f is the identity, *every* μ is a fixed point of T. For more general offspring distributions, whenever x is a fixed point of f, the Bernoulli distribution which puts mass x at 0 and $1 - x$ at 1 is a fixed point of T; for a game with Bernoulli terminal values, there are only two possible values of the outcome and we can interpret 0 as a win and 1 as a loss (from the perspective of the first player).

Suppose indeed that μ is a fixed point of the operator T. Consider a tree of depth $2n$ (given by the Galton–Watson tree truncated at level $2n$) with the terminal values drawn independently from μ. Then the distribution of the value at the root is also μ. More generally, consider the structure of the first k levels of the tree; the distribution of these first k levels is the same for any n (such that $k \leqslant 2n$).

As a consequence of this *consistency* over different values of n, we may let $n \to \infty$ and, applying Kolmogorov's extension theorem, obtain a distribution of the entire infinite Galton–Watson tree along with values attached to each node which obey the minimax recursions (min at even levels, max at odd levels).

This gives a *stationary recursive tree process* in the language of [1]. The relevant stationarity property is the following: condition on the structure of the first two levels of the tree, and write $v_1, \ldots, v_r$ for the level 2 nodes. Conditional on the structure of the first two levels, the structure of the subtrees rooted at $v_1, \ldots, v_r$, along with the values associated to the nodes of those subtrees, are given by r i.i.d. copies of the original tree process. (More precisely, we might describe the tree process as '2-periodic' rather than stationary, since even and odd levels differ; we can recover a genuinely stationary process by considering only even levels.)

For a more formal and more general set-up, see for example [1] or [9].

We have defined a joint distribution of the structure of the tree and the values associated to each node of the tree. Now the recursive tree process is said to be *endogenous* if the value associated to the root is measurable with respect to the structure of the tree. Note that for the same offspring

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

distribution, this endogeny property may hold for some fixed point distributions μ and not for others.

Being measurable with respect to the structure of the tree is equivalent to being approximable to any given degree of accuracy using the information only of some finite portion of the tree. That is, for any random variable X (in particular the root value), X is measurable with respect to the structure of the tree if, for any x and any $\varepsilon > 0$, there exists k such that with probability $1-\varepsilon$, the conditional probability of the event $\{X \leqslant x\}$, given the structure of the first k levels of the tree is in $[0, \varepsilon] \cup [1-\varepsilon, 1]$, where X denotes the value at the root.

For a more concrete interpretation, we can concentrate only on the case of finite trees, truncated at some level $2n$. Then the property in the previous paragraph can be reformulated to say that the value at the root can be approximated arbitrarily closely using information from the structure of some appropriate number of levels at the top of the tree, *uniformly* in the value of n.

Note that endogeny does *not* indicate that the value at the root is insensitive to arbitrary changes in the boundary conditions. In our case, that would be trivially false. Rather, for a given distribution of boundary conditions, endogeny expresses the property that, if the boundary is far away, the root is typically not much affected by the difference between various realizations drawn from that distribution. In particular, endogeny may hold for some boundary distributions and not for others, as is indeed the case for our model.

Consider in particular the Bernoulli ('win/loss') boundary conditions described above.

Theorem 3.1. *Let $x \in (0,1)$ be a fixed point of f, and consider the stationary recursive tree process with Bernoulli $(1-x)$ marginals for the values at even levels. The process is endogenous if and only if $f'(x) \leqslant 1$.*

So, approximately speaking, the endogenous processes with Bernoulli marginals correspond to the *stable* fixed points of the function f, which are those fixed points which do *not* appear as atoms in the distribution of the limiting random variable W in Theorem 1.2. (An exception may occur when the derivative of f at a fixed point is precisely 1; further, in the cases $x=0$ and $x=1$ the values are constant and the process is trivially endogenous.)

To prove Theorem 3.1, we use a characterization of endogeny in terms of uniqueness of bivariate distributions, introduced by Aldous and Bandyopadhyay in [1] and proved in somewhat more generality by Mach, Sturm and Swart [9]. See Section 5 for details.

For offspring distributions where f is the identity, any distribution μ gives rise to a recursive tree process. In particular, we can take μ to be the uniform distribution on $[0,1]$, as we did in previous sections. We have the following corollary of Theorem 3.1.

Corollary 3.2. *Let f be the identity. Then, for any μ, the recursive tree process with marginals μ for the values at even levels is endogenous.*

4. Proofs: convergence and scaling limits

First, we show how (1.3) follows from the recursive distributional equation (1.2). As at (1.2), let M and M_1, M_2, M_3 be i.i.d. draws from the offspring distribution, and $W_{2n-2}^{(i,j)}$ i.i.d. copies of the random variable W_{2n-2}, independent of M and $\{M_i, i \geqslant 1\}$.

Note that for any given i,

$$\begin{aligned}\mathbb{P}\Big(\max_{1\leqslant j\leqslant M_i} W_{2n-2}^{(i,j)} > x\Big) &= 1 - \mathbb{P}(W_{2n-2}^{(i,j)} \leqslant x \quad \text{for } j = 1, \dots, M_i) \\ &= 1 - \sum_m p_m \mathbb{P}(W_{2n-2} \leqslant x)^m \\ &= R(\mathbb{P}(W_{2n-2} \leqslant x)).\end{aligned}$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

So from (1.2) we have

$$\begin{aligned}\mathbb{P}(W_{2n} \leqslant x) &= \mathbb{P}\Big(\min_{1\leqslant i\leqslant M}\max_{1\leqslant j\leqslant M_i} W_{2n-2}^{(i,j)} \leqslant x\Big)\\ &= 1-\mathbb{P}\Big(\max_{1\leqslant j\leqslant M_i} W_{2n-2}^{(i,j)} > x \text{ for all } i=1,\dots,M\Big)\\ &= 1-\sum_m p_m\big[R(\mathbb{P}(W_{2n-2}\leqslant x))\big]^m\\ &= R(R(\mathbb{P}(W_{2n-2}\leqslant x)))\\ &= f(\mathbb{P}(W_{2n-2}\leqslant x)),\end{aligned}$$

giving (1.3) as desired.

Proof of Theorem 1.2. From the previous line and the monotonicity of f we see that

$$\lim_{n\to\infty}\mathbb{P}(W_{2n}\leqslant x) = \lim_{n\to\infty} f^n(x)$$

exists for all x, and therefore W_{2n} indeed converges in distribution as $n\to\infty$, to a limit W with the distribution function $F_W(x)=\lim_{n\to\infty} f^n(x)$.

Part (a) is immediate from (1.3). For part (b), note that since f is analytic in $(0,1)$ and f is not the identity function, the set of fixed points of f cannot have an accumulation point in $(0,1)$. Therefore, this set of fixed points of f defines a partition of the interval $(0,1)$ into a disjoint union of open intervals plus the set of fixed points, each of which is an endpoint of exactly two intervals from the partition. Since f is monotone and continuous, $F_W(x)=\lim_{n\to\infty} f^n(x)$ is constant on those intervals; therefore W can have atoms only at fixed points of f.

Suppose $q\in(0,1)$ is such a fixed point. Then

$$\mathbb{P}(W=x) = \lim_{\varepsilon\to 0}\mathbb{P}(q-\varepsilon < W\leqslant x+\varepsilon) = \lim_{\varepsilon\to 0}\lim_{n\to\infty} f^n(q+\varepsilon) - \lim_{\varepsilon\to 0}\lim_{n\to\infty} f^n(q-\varepsilon). \tag{4.1}$$

Since f is monotone and continuous, the quantity above is equal to 0 precisely if and only if the fixed point q is stable. Hence indeed W has an atom at q precisely if q is unstable from at least one side. As commented immediately after Theorem 1.2, the right-hand side of (4.1) is equal to $q_+ - q_-$, as required.

The cases where $q=0$ or $q=1$ follow in a similar way. □

The rest of this section is devoted to the proof of Theorem 1.3.

4.1 Proof of Theorem 1.3(a): $1 < f'(q) < \infty$

Firstly we assume that q is the unique fixed point of f in $(0,1)$ and that it is unstable from both sides. In the second part of the proof we show how Lemma 4.1 below implies the general case.

Suppose $\xi = f'(q) > 1$. An example of this case is presented in Figure 2(a), where $p_1 = 0.45$ and $p_3 = 0.55$.

We will prove the following result.

Lemma 4.1. *Consider the recursion (1.3) and assume that q is the unique fixed point of f in $(0,1)$ and that it is unstable from both sides. Set $\xi = f'(q)$. If $\xi > 1$, then*

$$\xi^n(W_{2n}-q) \xrightarrow{d} V \quad \text{as } n\to\infty,$$

where the distribution function F_V of V is continuous and satisfies

$$F_V(x) = f(F_V(x/\xi)), \quad x\in\mathbb{R}.$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Lemma 4.1 extends the result of Ali Khan, Devroye and Neininger [2] to the case of random trees. Note that Lemma 4.1 corresponds directly to part (a) of Theorem 1.3 for f having a unique fixed point in $(0, 1)$ which is unstable, as then $q_- = 0$ and $q_+ = 1$.

Proof of Lemma 4.1. We follow the lines of [2] but in our case the analysis is slightly more complicated because of the more general form that f can admit. We first prove that there exists a pointwise limit of distribution functions of $\xi^n(W_{2n} - q)$, which is not identical to 0 or 1, and then show that it is continuous, which completes the proof. Define

$$g_n(x) = \mathbb{P}(\xi^n(W_{2n} - q) \leqslant x), \quad x \in \mathbb{R}.$$

Therefore, for each x for sufficiently large n (such that $0 \leqslant q + x/\xi^n \leqslant 1$),

$$g_n(x) = \mathbb{P}\left(W_{2n} \leqslant q + \frac{x}{\xi^n}\right) = f^n\left(q + \frac{x}{\xi^n}\right).$$

Note that $g_n(0) = q$ for all n. We need some local uniform bound for g_n around $x = 0$. This will be supplied by the following lemma.

Lemma 4.2. *Under the assumptions of Lemma 4.1, let k be the smallest number larger than 1 such that $f^{(k)}(q) \neq 0$. Denote $h_1(x) = q + x$ and $h_2(x) = q + x + cx^k$ for $x \in \mathbb{R}$. Then there exist c and an $\varepsilon > 0$ such that, for all n and $|x| \leqslant \varepsilon$, either $h_1(x) \leqslant g_n(x) \leqslant h_2(x)$ or $h_2(x) \leqslant g_n(x) \leqslant h_1(x)$.*

Note that such a number k exists since we assumed that f is not the identity function and f is analytic at q.

Proof of Lemma 4.2. Take any c such that c has the same sign as $f^{(k)}(q)$ and

$$|c| > \left|\frac{f^{(k)}(q)}{k!\xi(\xi^{k-1} - 1)}\right|.$$

From analyticity of f, $f^{(k)}(x)$ does not change sign on some neighbourhood of q.

For simplicity assume that k is even and $f^{(k)}(q) > 0$. We would generally need to consider four cases depending on the parity of k and the sign of $f^{(k)}(q)$. For the other three cases the steps of the proof of the lemma are identical modulo the change of sign in the inequalities.

The proof is by induction on n and makes use of Taylor's formula up to order k. For $n = 0$ the assertion is true, as we have

$$h_1(x) = q + x = g_0(x) \leqslant h_2(x).$$

Note that the above holds for all ε, thus we will choose ε later. Assume now that

$$h_1(x) \leqslant g_{n-1}(x) \leqslant h_2(x)$$

for some $n - 1 \geqslant 0$, $|x| \leqslant \varepsilon$ and $\varepsilon > 0$. Since $|x/\xi| \leqslant \varepsilon$, as $|x| \leqslant \varepsilon$ and $\xi > 1$, and f is increasing, we have

$$g_n(x) = f^n\left(q + \frac{x}{\xi^n}\right) = f\left(f^{n-1}\left(q + \frac{x/\xi}{\xi^{n-1}}\right)\right) = f(g_{n-1}(x/\xi)) \geqslant f(h_1(x/\xi)),$$

and analogously

$$g_n(x) \leqslant f(h_2(x/\xi)).$$

The induction proof will be completed if we can show that for some $\varepsilon > 0$, for $|x| \leqslant \varepsilon$,

$$h_1(x) \leqslant f(h_1(x/\xi)), \quad f(h_2(x/\xi)) \leqslant h_2(x). \tag{4.2}$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Taking the Taylor expansion of f around q at points $q + x/\xi$ and $q + x/\xi + c(x/\xi)^k$, we obtain

$$f(h_1(x/\xi)) = q + x + \frac{1}{k!}\frac{1}{\xi^k}(f^{(k)}(q))x^k + o(x^k),$$
$$f(h_2(x/\xi)) = q + x + \frac{1}{k!}\frac{1}{\xi^k}(f^{(k)}(q) + \xi k!c)x^k + o(x^k).$$

Since

$$0 < \frac{1}{k!}\frac{1}{\xi^k}(f^{(k)}(q) + \xi k!c) = \frac{f^{(k)}(q)}{k!\xi(\xi^{k-1}-1)}\frac{\xi^{k-1}-1}{\xi^{k-1}} + c\frac{1}{\xi^{k-1}} < c$$

and by assumption $f^{(k)}(q) > 0$, we are therefore able to pick $\varepsilon > 0$ such that (4.2) holds for $|x| \leqslant \varepsilon$. This completes the proof of Lemma 4.2. □

Note that if we show that for some g, $g_n(x) \to g(x)$ for all x, then the above lemma will imply that $g(x)$ is continuous and differentiable at $x = 0$ with $g'(0) = 1$. We now claim that for each x, $(g_n(x))$ is a monotone sequence for $n > n_x$. This is implied by the following lemma.

Lemma 4.3. *Under the assumptions of Lemma 4.1, for each M there exists n_M such that for $|x| \leqslant M$, $(g_n(x))$ is a monotone sequence for $n \geqslant n_M$.*

Proof of Lemma 4.3. As in the proof of Lemma 4.2, we consider the case where k is even and $f^{(k)}(q) > 0$; the other cases are identical. Using Taylor expansion up to order k, there exists $\varepsilon > 0$ such that, for $|y| \leqslant \varepsilon$,

$$f(q + y) \geqslant f(q) + f'(q)y \tag{4.3}$$

(recall that $f^{(i)}(q) = 0$ for $1 < i < k$). Now let $n_M = \lceil \log_\xi (M/\varepsilon) \rceil$ and note that for any $|x| \leqslant M$ and $n \geqslant n_M$, $|x/\xi^n| < \varepsilon$ and therefore by (4.3),

$$f\left(q + \frac{x}{\xi^n}\right) \geqslant f(q) + f'(q)\frac{x}{\xi^n} = q + \frac{x}{\xi^{n-1}}.$$

Finally, since f^{n-1} is monotone increasing,

$$g_n(x) = f^n\left(q + \frac{x}{\xi^n}\right) = f^{n-1}\left(f\left(q + \frac{x}{\xi^n}\right)\right) \geqslant f^{n-1}\left(q + \frac{x}{\xi^{n-1}}\right) = g_{n-1}(x).$$

This proves the claim. □

Since $g_n(x) \leqslant 1$, by Lemma 4.3 $g_n(x)$ converges for all x; we denote the limit by $g(x)$. The continuity of f and the fact that $g_n(x) = f(g_{n-1}(x/\xi))$ imply that

$$g(x) = f(g(x/\xi)). \tag{4.4}$$

Therefore, from the continuity of f and the monotonicity of g, $\lim_{x\to-\infty} g(x)$ and $\lim_{x\to\infty} g(x)$ are fixed points of f. Using the fact that $\{0, q, 1\}$ are the only fixed points of f, g is non-decreasing, $g(0) = q$ and $g'(0) = 1$, we deduce that $\lim_{x\to-\infty} g(x) = 0$ and $\lim_{x\to\infty} g(x) = 1$. When we show that g is continuous at all x, it will then imply that $F_V = g$.

We apply the following strategy to show that g is continuous: we showed that $g(x)$ is continuous at 0 and now we show separately that it is continuous on some $(-\varepsilon, 0)$ and on some $(0, \varepsilon)$. The identity (4.4), together with the continuity of f, then implies that g is continuous on all of $\mathbb{R}$ (since $\xi > 1$).

We still work under the assumption that $f^{(k)}(q) > 0$, where $k \geqslant 2$ is such that $f^{(r)}(q) = 0$ for $1 < r < k$ and $f^{(k)}(q) \neq 0$, and that k is even (if k is odd then reasoning in the two cases below

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

should be swapped). Note that to prove that g is continuous on some interval I, it is sufficient to show that

$$\sup_{y\in I}\sup_{n\geqslant 0} g_n'(y) < \infty. \tag{4.5}$$

By the chain rule we obtain

$$g_n'(x) = \frac{1}{\xi^n}\prod_{i=0}^{n-1} f'\left(f^i\left(q+\frac{x}{\xi^n}\right)\right). \tag{4.6}$$

We consider first $g_n'(y)$ for $y<0$. Since $f^{(k)}(q)>0$, there exists an $\varepsilon>0$ such that $f'(q+y)<\xi$ for $y\in(-\varepsilon,0)$. Since $f(q)=q$ and $\xi>1$, this implies that for $i<n$

$$q > f^i\left(q+\frac{y}{\xi^n}\right) > q+\xi^i\frac{y}{\xi^n} > q-\varepsilon,$$

hence

$$f'\left(f^i\left(q+\frac{y}{\xi^n}\right)\right) < \xi.$$

By (4.6) we conclude that $g_n'(y)<1$ for all n and $y\in(-\varepsilon,0)$. This implies that (4.5) holds with $I=(-\varepsilon,0)$, hence g is continuous on $(-\varepsilon,0)$.

Now we turn to the case of $y>0$. The function f is non-decreasing, and $\xi>1$; hence, for all $0<i<n$ and all $0\leqslant y\leqslant x$,

$$q \leqslant f^i\left(q+\frac{y}{\xi^n}\right) \leqslant f^i\left(q+\frac{x}{\xi^n}\right). \tag{4.7}$$

Note also that

$$f^i\left(q+\frac{x}{\xi^n}\right) \leqslant f^i\left(q+\frac{x}{\xi^i}\right) = g_i(x). \tag{4.8}$$

Now by the assumption $f^{(k)}>0$ there exists an $\varepsilon>0$ such that $f'(q+x)$ is strictly increasing for $x\in(0,\varepsilon)$. By the continuity of g at 0, there exists $\gamma>0$ such that $g(x)<q+\varepsilon$ for $x\in(0,\gamma)$. By Lemma 4.3, there exists n_γ such that, for $0<x\leqslant\gamma$, $(g_i(x))$ is a monotone sequence (an increasing one, since we assume $f^{(k)}(q)>0$) for $i>n_\gamma$. Therefore, for $i\geqslant n_\gamma$, and for $0<x<\gamma$,

$$g_i(x) \leqslant g(x) < q+\varepsilon. \tag{4.9}$$

On the other hand, for $i<n_\gamma\wedge n$, since $f(q+x)>q+x$ for $x\in(0,1-q)$,

$$f^i\left(q+\frac{x}{\xi^n}\right) \leqslant f^{n_\gamma}\left(q+\frac{x}{\xi^n}\right) \leqslant f^{n_\gamma}(q+x). \tag{4.10}$$

f^{n_γ} is continuous and non-decreasing, hence we may pick $\tilde\gamma>0$ such that, for $0<x<\tilde\gamma$,

$$f^{n_\gamma}(q+x) < q+\varepsilon. \tag{4.11}$$

Finally, combining (4.7)–(4.11), we obtain that for all $0<i<n$ and for all $0\leqslant y\leqslant x\leqslant\gamma\wedge\tilde\gamma$,

$$q \leqslant f^i\left(q+\frac{y}{\xi^n}\right) \leqslant f^i\left(q+\frac{x}{\xi^n}\right) \leqslant q+\varepsilon. \tag{4.12}$$

Recall that ε was chosen to be such that f' is strictly increasing on $(q,q+\varepsilon)$. Combining this with (4.6) and (4.12), we obtain that

$$g_n'(y) \leqslant g_n'(x). \tag{4.13}$$

We are now going to use (4.13) to show that (4.5) holds for $I=(0,\varepsilon)$ with $\varepsilon=\frac{1}{2}(\gamma\wedge\tilde\gamma)$. For each $z\in(\varepsilon,2\varepsilon)$ and all $n\geqslant 0$, by (4.13) we have

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

$$\sup_{y\in(0,\varepsilon)} g_n'(y) \leqslant g_n'(z).$$

Therefore,

$$\sup_{n\geqslant 0}\sup_{y\in(0,\varepsilon)} g_n'(y) \leqslant \sup_{n\geqslant 0}\frac{1}{\varepsilon}\int_{\varepsilon}^{2\varepsilon} g_n'(z)\,\mathrm{d}z \leqslant \sup_{n\geqslant 0}\frac{1}{\varepsilon}(g_n(2\varepsilon)-g_n(\varepsilon) \leqslant \frac{1}{\varepsilon},$$

where the last inequality follows since each g_n is a distribution function. This proves that $g(y)$ is continuous on $(0,\varepsilon)$. This completes the proof of Lemma 4.1. □

Multiple atoms. In the previous section we found the correct order of fluctuations when f had a single fixed point in the interval $(0,1)$. When f has more than one fixed point in the interval $(0,1)$, we cannot simply consider the quantity $W_{2n}-q$, since the limiting distribution has multiple atoms, but it turns out that we can condition on W_{2n} being close enough to one of the atoms and straightforwardly apply Lemma 4.1. Note that the set of fixed points of f cannot have an accumulation point in the interval $(0,1)$. To see this, recall that f is a composition of functions analytic in $(0,1)$. Therefore, $f(x)-x$ is also analytic in $(0,1)$ and we justify the claim using the fact that zeros of an analytic function not identical to 0 cannot have any accumulation points in the domain in which the function is analytic.

The case of multiple atoms of V is summarized in the following lemma.

Lemma 4.4. *Consider the recursion* (1.3) *and assume that* q_-, q, q_+ *are fixed points of* f *satisfying the following conditions:*

- $q\in(0,1)$ *is unstable and* $f'(q)>1$,
- $q_-<q<q_+$,
- q *is the only unstable point from at least one side point of* f *in the interval* (q_-,q_+).

Then

$$\mathcal{L}(\xi^n(W_{2n}-q)\mid W_{2n}\in[q_-,q_+])\to\mathcal{L}(V),$$

where $\xi=f'(q)$ *and* V *is a random variable with a continuous distribution function.*

Note first that since $f'(q)>1$, the definitions of q_- and q_+ coincide with those given in (1.4).

Proof of Lemma 4.4. Fix $x\in[0,1]$. Then

$$\begin{aligned}
&\mathbb{P}\left(\frac{W_{2n}-q_-}{q_+-q_-}\leqslant x\,\middle|\, W_{2n}\in[q_-,q_+]\right)\\
&\quad=\frac{\mathbb{P}(q_-\leqslant W_{2n}\leqslant x(q_+-q_-)+q_-)}{\mathbb{P}(q_-\leqslant W_{2n}\leqslant q_+)}\\
&\quad=\frac{f(\mathbb{P}(W_{2n-2}\leqslant x(q_+-q_-)+q_-)-q_-}{q_+-q_-}\\
&\quad=\frac{f(\mathbb{P}((W_{2n-2}-q_-)/(q_+-q_-)\leqslant x))-q_-}{q_+-q_-}\\
&\quad=\tilde{f}\left(\mathbb{P}\left(\frac{W_{2n-2}-q_-}{q_+-q_-}\leqslant x\,\middle|\, W_{2n-2}\in[q_-,q_+]\right)\right),
\end{aligned}\tag{4.14}$$

where

$$\tilde{f}(x)=\frac{f(x(q_+-q_-)+q_-)-q_-}{q_+-q_-}.$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Furthermore, $\tilde{f}(x)$ is a continuous bijective mapping from $[0, 1]$ to $[0, 1]$ with a single fixed point $\tilde{q} = (q - q_-)/(q_+ - q_-)$ in $(0, 1)$ and

$$\begin{aligned} \tilde{f}'(\tilde{q}) &= f'(q) = \xi, \\ \tilde{f}^{(k)}(\tilde{q}) &= f^{(k)}(q)(q_+ - q_-)^{k-1}. \end{aligned} \tag{4.15}$$

Consider a sequence of random variables $(\tilde{W}_{2n})_{n=0}^{\infty}$ such that

$$\tilde{W}_{2n} \stackrel{d}{=} \left(\frac{W_{2n} - q_-}{q_+ - q_-} \;\middle|\; W_{2n} \in [q_-, q_+] \right).$$

We check that

$$\tilde{W}_0 \sim U(0, 1), \quad \mathbb{P}(\tilde{W}_{2n} \leqslant x) = \tilde{f}(\mathbb{P}(\tilde{W}_{2n-2} \leqslant x).$$

Combining this with (4.15), we may apply Lemma 4.1 to $(\tilde{W}_{2n})_{n=0}^{\infty}$ to conclude that

$$\mathcal{L}(\xi^n (W_{2n} - q) \mid W_{2n} \in [q_-, q_+]) \to \mathcal{L}((q_+ - q_-)V),$$

where W is a random variable with a continuous distribution function $\tilde{g}$ that satisfies $\tilde{g}(x) = \tilde{f}(\tilde{g}(x))$. Finally, we note that the distribution function of $(q_+ - q_-)V$ is also continuous, which completes the proof. □

Boundary fixed points. To finish the proof of part (a) of Theorem 1.3 we need to consider the case when one of q_-, q_+ is equal to q. This may happen either if q is at the boundary (*i.e.* $q \in \{0, 1\}$) or when $q \in (0, 1)$, but q is stable from one side. These cases can be treated simultaneously by repeating the reasoning from the proofs of Lemmas 4.1 and 4.4. Note that the limiting distribution V is now concentrated on either the positive or negative half-line. □

4.2 Proof of Theorem 1.3(b): $f'(q) = 1$

We have already described the fluctuations of W_{2n} when we know that it converges to some fixed point q of f with $f'(q) \in (1, \infty)$. If the point was unstable from both sides, we obtained a two-sided continuous limiting distribution.

If q is a fixed point of f such that $f'(q) = 1$, it may be unstable, stable, or unstable from one side and stable from the other. In this case it is more convenient to consider each side of q separately. For simplicity, we state and prove a lemma for the case where q is unstable from the right and then comment on the general case.

Note that the set of fixed points does not have an accumulation point in $(0, 1)$, but it is not known whether this behaviour may be exhibited at the boundary, hence the additional assumption in the lemma.

Lemma 4.5. *Consider the recursion (1.3) and assume that q is a fixed point of f, that it is unstable from the right and let $q_+ = \inf\{x : x > q, x = f(x)$. Suppose that $f'(q) = 1$ and k is such that $f^{(r)}(q) = 0$ for $1 < r < k$ and $f^{(k)}(q) \neq 0$. If $q_+ \neq q$, then*

$$\mathcal{L}(n^{1/(k-1)}(W_{2n} - q) \mid W_{2n} \in [q, q_+]) \to \delta_a,$$

where

$$a = \left(\frac{k(k-2)!}{f^{(k)}(q)} \right)^{1/(k-1)}.$$

Figure 3(a) gives an example where Lemma 4.5 applies.

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Proof of Lemma 4.5. We are going to show that the distribution function of

$$n^{1/(k-1)}\frac{W_{2n}-q}{q_+-q}$$

conditioned on the event $W_{2n}\in[q,q_+]$ converges to some limit as n tends to infinity. Define

$$g_n(x):=\mathbb{P}\left(n^{1/(k-1)}\frac{W_{2n}-q}{q_+-q}\leqslant x\;\middle|\;W_{2n}\in[q,q_+]\right). \tag{4.16}$$

By calculations similar to those of (4.14) in the proof of Lemma 4.4, we obtain that for $x\in[0,1]$,

$$\mathbb{P}\left(\frac{W_{2n}-q}{q_+-q}\leqslant x\;\middle|\;W_{2n}\in[q,q_+]\right)=\tilde{f}\left(\mathbb{P}\left(\frac{W_{2n}-q}{q_+-q}\leqslant x\;\middle|\;W_{2n-2}\in[q,q_+]\right)\right),$$

where

$$\tilde{f}(x)=\frac{f(x(q_+-q)+q)-q}{q_+-q}.$$

Note that

$$\begin{aligned}\tilde{f}(0)&=0,\\ \tilde{f}(1)&=1,\\ \tilde{f}'(0)&=f'(q)=1,\\ \tilde{f}^{(i)}(0)&=(q_+-q)^{i-1}f^{(i)}(q),\end{aligned}$$

and $\tilde{f}$ has no fixed points in $(0,1)$. Since for every $x>0$, for sufficiently large n,

$$\frac{x}{n^{1/k-1}}\in[0,1],$$

for such n we have

$$\begin{aligned}g_n(x)&=\mathbb{P}\left(\frac{W_{2n}-q}{q_+-q}\leqslant xn^{-1/(k-1)}\;\middle|\;W_{2n}\in[q,q_+]\right)\\ &=\tilde{f}^n\left(\mathbb{P}\left(\frac{W_0-q}{q_+-q}\leqslant xn^{-1/(k-1)}\;\middle|\;W_0\in[q,q_+]\right)\right)\\ &=\tilde{f}^n(xn^{-1/(k-1)}).\end{aligned} \tag{4.17}$$

The proof consists of two parts.

(I) We show that for each $x<a$, for sufficiently large n, $(g_n(x))$ forms a decreasing sequence, and for each $x>a$, for sufficiently large n, $(g_n(x))$ forms an increasing sequence.
(II) We show that for $x<a$, $g_n(x)\to 0$, and for $x>a$, $g_n(x)\to 1$.

Part I. Fix $x\neq 0$. Using Taylor's expansion, we may expand $\tilde{f}(x)$ as follows:

$$\tilde{f}(x)=\tilde{f}(0)+x+\frac{\tilde{f}^{(k)}(0)}{k!}x^k+r_k(x)x^k,$$

where $\lim_{x\to 0}r_k(x)=0$. Therefore

$$\tilde{f}\left(\frac{x}{n^{1/(k-1)}}\right)<\frac{x}{(n-1)^{1/(k-1)}} \tag{4.18}$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

is equivalent to

$$\frac{x}{n^{1/(k-1)}}+\frac{\tilde{f}^{(k)}(0)}{k!}\left(\frac{x}{n^{1/(k-1)}}\right)^k+r_k\left(\frac{x}{n^{1/(k-1)}}\right)\left(\frac{x}{n^{1/(k-1)}}\right)^k<\frac{x}{(n-1)^{1/(k-1)}},$$

and to

$$\frac{\tilde{f}^{(k)}(0)}{k!}x^{k-1}+r_k\left(\frac{x}{n^{1/(k-1)}}\right)x^{k-1}<n\left(\left(\frac{n}{n-1}\right)^{1/(k-1)}-1\right).$$

Letting $n\to\infty$, the right-hand side of the last formula converges to $1/(k-1)$, whereas the left-hand side converges to

$$\frac{\tilde{f}^{(k)}(0)}{k!}x^{k-1}.$$

Therefore, the last inequality is satisfied for large n if

$$x<\left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{1/(k-1)}\frac{1}{q_+-q},$$

and similarly

$$\tilde{f}\left(\frac{x}{n^{1/(k-1)}}\right)>\frac{x}{(n-1)^{1/(k-1)}} \tag{4.19}$$

for large n if

$$x>\left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{1/(k-1)}\frac{1}{q_+-q}.$$

This yields the claim, as $\tilde{f}^{n-1}$ is a strictly increasing function, hence recalling (4.17), inequality (4.18) is equivalent to

$$g_n(x)=\tilde{f}^n\left(\frac{x}{n^{1/(k-1)}}\right)<\tilde{f}^{n-1}\left(\frac{x}{(n-1)^{1/(k-1)}}\right)=g_{n-1}(x),$$

and inequality (4.19) is equivalent to

$$g_n(x)=\tilde{f}^n\left(\frac{x}{n^{1/(k-1)}}\right)>\tilde{f}^{n-1}\left(\frac{x}{(n-1)^{1/(k-1)}}\right)=g_{n-1}(x).$$

This completes the proof of the claim.

Part II. Since each g_n is a distribution function, and by Part I above for each x, $(g_n(x))$ is a monotone sequence for large n (decreasing for $x<a$ and increasing for $x>a$), hence the limit $g(x)=\lim_{n\to\infty}g_n(x)$ exists for all x. To show that for $x<a$, $g(x)=0$, assume that for some $0<x<a$, $g(x)=\varepsilon>0$. This implies that

$$\tilde{f}^n\left(\frac{x}{n^{1/(k-1)}}\right)\geqslant\varepsilon>0 \tag{4.20}$$

for large n. Take $y\in\mathbb{R}$, $l\in\mathbb{N}$ such that

$$y=x\left(\frac{l}{l-1}\right)^{1/(k-1)}<a.$$

Note also that $g(y)\leqslant 1$, but since $(g_n(y))$ is a strictly decreasing sequence, the inequality is in fact sharp, thus

$$\lim_{n\to\infty}g_n(y)<1. \tag{4.21}$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

On the other hand,

$$\begin{aligned}\lim_{n\to\infty} g_n(y) &= \lim_{n\to\infty} g_{nl}(y) = \lim_{n\to\infty} \tilde{f}^{nl}\left(\frac{y}{(nl)^{1/(k-1)}}\right)\\ &= \lim_{n\to\infty} \tilde{f}^{n} \circ \tilde{f}^{n(l-1)}\left(\frac{y((l-1)/l)^{1/(k-1)}}{(n(l-1))^{1/(k-1)}}\right)\\ &= \lim_{n\to\infty} \tilde{f}^{n} \circ \tilde{f}^{n(l-1)}\left(\frac{x}{(n(l-1))^{1/(k-1)}}\right),\end{aligned} \tag{4.22}$$

and by (4.20),

$$\lim_{n\to\infty} \tilde{f}^{n} \circ \tilde{f}^{n(l-1)}\left(\frac{x}{(n(l-1))^{1/(k-1)}}\right) \geqslant \lim_{n\to\infty} \tilde{f}^{n}(\varepsilon) = 1,$$

as 1 is the only stable fixed point of $\tilde{f}$ in the interval $[0, 1]$. But this contradicts (4.21) and thus $g(x) = 0$.

Similarly, to show that for $x > a$, $g(x) = 1$, fix any such x and take $y \in \mathbb{R}$, $l \in \mathbb{N}$ such that

$$y = \left(\frac{l-1}{l}\right)^{1/(k-1)} x > a.$$

By calculations similar to (4.22),

$$\begin{aligned}\lim_{n\to\infty} g_n(x) &= \lim_{n\to\infty} g_{nl}(x) = \lim_{n\to\infty} \tilde{f}^{nl}\left(\frac{x}{(nl)^{1/(k-1)}}\right)\\ &= \lim_{n\to\infty} \tilde{f}^{n} \circ \tilde{f}^{n(l-1)}\left(\frac{y}{(n(l-1))^{1/(k-1)}}\right)\\ &= \lim_{n\to\infty} \tilde{f}^{n} \circ g_{n(l-1)}(y),\end{aligned}$$

but since $(g_n(y))$ is a strictly increasing sequence for large n, for these n, $g_n(y) \geqslant \delta$ for some $\delta > 0$, hence

$$\lim_{n\to\infty} \tilde{f}^{n} \circ g_{n(l-1)}(y) \geqslant \lim_{n\to\infty} \tilde{f}^{n}(\delta) = 1,$$

as again, 1 is the only stable fixed point of $\tilde{f}$ in $[0, 1]$.

Recall now the definition of $g_n(x)$ (4.16). Parts I and II prove that

$$\mathcal{L}\left(n^{1/(k-1)}\frac{W_{2n}-q}{q_+-q} \;\middle|\; W_{2n} \in [q, q_+]\right) \to \delta_a,$$

and thus

$$\mathcal{L}(n^{1/(k-1)}(W_{2n}-q) \mid W_{2n} \in [q, q_+]) \to \delta_{a(q_+-q)},$$

where

$$a(q_+-q) = \left(\frac{k(k-2)!}{f^{(k)}(q)}\right)^{1/(k-1)}.$$

This completes the proof of Lemma 4.5. □

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

To finish the proof of part (b) of Theorem 1.3 we apply Lemma 4.5 and its counterpart for points unstable from the left to $[q, q_+]$ and $[q_-, q]$ respectively. Checking that for each n

$$P(W_{2n} \in [q, q_+] \mid W_{2n} \in [q_-, q_+]) = \frac{q_+ - q}{q_+ - q_-},$$
$$P(W_{2n} \in [q_-, q] \mid W_{2n} \in [q_-, q_+]) = \frac{q - q_-}{q_+ - q_-},$$

shows that the masses in the formulation of the theorem are chosen appropriately, hence completes the proof. □

4.3 Proof of Theorem 1.3(c): $f'(q) = \infty$

Note first that $f'(q) = \infty$ can happen only at $q \in \{0, 1\}$.

We start by describing the behaviour of f near 0. The first step is supplied by the following technical lemma.

Lemma 4.6. *There exist functions $H(x)$ and $b(x)$ defined on $(-1, 1)$ such that $f(x) \sim H(x)$ as $x \to 0$ and*

$$H'(x) \sim \left(\sum_{n=1}^{\infty} n p_n b(x)^{n-1}\right) K p_K x^{K-1},$$

where $1 - b(x) \sim p_K x^K$ as $x \to 0$.

Proof of Lemma 4.6. By simple calculations,

$$\begin{aligned} f(x) = R(R(x)) &= \frac{R(R(x))}{1 - R(x)}(1 - R(x)) \\ &= \frac{1 - \sum_{n=1}^{\infty} p_n R(x)^n}{1 - R(x)} \sum_{n=1}^{\infty} p_n x^n \\ &= \frac{\sum_{n=1}^{\infty} p_n (1 - R(x)^n)}{1 - R(x)} \sum_{n=1}^{\infty} p_n x^n \\ &= \sum_{n=1}^{\infty} \left(p_n \sum_{i=0}^{n-1} R(x)^i\right) \sum_{n=1}^{\infty} p_n x^n \\ &= \sum_{i=0}^{\infty} \left(R(x)^i \sum_{n=i+1}^{\infty} p_n\right) \sum_{n=1}^{\infty} p_n x^n \\ &= \sum_{i=0}^{\infty} [R(x)^i \mathbb{P}(M > i)] \sum_{n=1}^{\infty} p_n x^n, \end{aligned} \tag{4.23}$$

where M is a random variable with law $\mathbb{P}(M = i) = p_i$. Furthermore, recalling that

$$K = \min\{i : P_i \neq 0\},$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

we have

$$\begin{aligned}\frac{R(x)}{1-x} &= \frac{\sum_{n=1}^{\infty} p_n(1-x^n)}{1-x} \\ &= \sum_{n=1}^{\infty} p_n \sum_{i=0}^{n-1} x^i \\ &= \sum_{i=0}^{\infty} x^i \sum_{n=i+1}^{\infty} p_n \\ &= \sum_{i=0}^{\infty} x^i \mathbb{P}(M > i) = 1 + x + \cdots + x^{K-1} + x^K \sum_{i=K}^{\infty} x^{i-K} \mathbb{P}(M > i). \end{aligned} \tag{4.24}$$

Therefore, substituting (4.24) into (4.23),

$$f(x) = \sum_{i=0}^{\infty} [(1-x)(1 + x + \cdots + x^{K-1} + x^K h(x))]^i \mathbb{P}(M > i) \sum_{n=1}^{\infty} p_n x^n, \tag{4.25}$$

where

$$h(x) = \sum_{i=K}^{\infty} x^{i-K} \mathbb{P}(M > i) \to \mathbb{P}(M > K) = 1 - p_K \text{ as } x \to 0. \tag{4.26}$$

Observe that $h'(x) \to \mathbb{P}(M > K+1)$ as $x \to 0$. Now, for any $b < 1$,

$$\sum_{i=0}^{\infty} b^i \mathbb{P}(M > i) = \sum_{n=1}^{\infty} p_n \sum_{i=0}^{n-1} b^i = \frac{1}{1-b}\left(1 - \sum_{n=1}^{\infty} p_n b^n\right),$$

and thus, setting

$$b(x) = (1-x)(1 + x + \cdots + x^{K-1} + x^K h(x)) = 1 - x^K + x^K h(x) - x^{K+1} h(x),$$

from (4.25) we obtain

$$f(x) = \frac{1}{1-b(x)}\left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right) \sum_{n=1}^{\infty} p_n x^n. \tag{4.27}$$

Observe that

$$1 - b(x) = x^K(1 - h(x)) + x^{K+1} h(x),$$

hence by (4.26),

$$1 - b(x) \sim p_K x^K \text{ as } x \to 0.$$

Moreover, from (4.27),

$$\begin{aligned} f(x) &= \frac{1}{x^K(1-h(x)) + x^{K+1}h(x)}\left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right) \sum_{n=1}^{\infty} p_n x^n \\ &= \frac{1}{1 - h(x) + xh(x)}\left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right) \sum_{n=K}^{\infty} p_n x^{n-K}. \end{aligned}$$

Note that as $x \to 0$, the first fraction on the right-hand side converges to $1/p_K$, and the final sum converges to p_K. Thus

$$f(x) \sim \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right) \text{ as } x \to 0.$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Therefore, denoting

$$H(x) := \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right),$$

we have

$$f(x) \sim H(x) \tag{4.28}$$

and, recalling that $\lim_{x\to 0} h'(x) = \mathbb{P}(M > K+1)$,

$$H'(x) = \left(1 - \sum_{n=1}^{\infty} p_n b(x)^n\right)' \sim \left(\sum_{n=1}^{\infty} n p_n b(x)^{n-1}\right) K p_K x^{K-1} \text{ as } x \to 0^+,$$

which completes the proof of the lemma. □

Equipped with the relation from Lemma 4.6 we may now connect f with the underlying offspring distribution via *Karamata's Tauberian theorem for power series* (a proof may be found in [3], for example). First recall the theorem.

Theorem 4.7. (Karamata's Tauberian theorem). *If the coefficients satisfy $a_n \geqslant 0$ and the power series $A(s) = \sum_{n=0}^{\infty} a_n s^n$ converges for $s \in [0,1)$, then for $c, \rho \geqslant 0$ the following are equivalent:*

$$\sum_{k=0}^{n} a_k \sim c n^{\rho} \quad \text{as } n \to \infty$$

and

$$A(s) \sim \frac{c\Gamma(1+\rho)}{(1-s)^{\rho}} \quad \text{as } s \uparrow 1.$$

Recall the assumption 1.5 of Theorem 1.3: for some $\rho \in (0,1)$,

$$\mathbb{E}(M \mathbb{I}_{M \leqslant n}) = \sum_{k=1}^{n} k p_k \sim c n^{\rho}.$$

By Theorem 4.7 applied to $a_k = k p_k$ and (1.5), we obtain that as we let $x \to 0$ (which implies $b(x) \to 1$),

$$\frac{1}{K p_K x^{K-1}} H'(x) \sim \sum_{n=1}^{\infty} n p_n b(x)^{n-1} \sim \frac{c\Gamma(1+\rho)}{(1-b(x))^{\rho}} \sim \frac{c\Gamma(1+\rho)}{p_K^{\rho}} \frac{1}{x^{K\rho}}.$$

Therefore,

$$H(t) = \int_0^t H'(x)\,\mathrm{d}x \sim \int_0^t \frac{c\Gamma(1+\rho)}{p_K^{\rho}} \frac{1}{x^{K\rho}} K p_K x^{K-1}\,\mathrm{d}x = \frac{c\Gamma(1+\rho) p_K^{1-\rho}}{1-\rho} t^{K-K\rho} \text{ as } t \to 0, \tag{4.29}$$

hence, by (4.28) and (4.29),

$$f(t) \sim \frac{c\Gamma(1+\rho) p_K^{1-\rho}}{1-\rho} t^{K(1-\rho)} \text{ as } t \to 0.$$

This implies that for $f'(0) = \infty$ to hold it is necessary that $K < 1/(1-\rho)$.

To provide the criterion for $q = 1$, we are interested in the behaviour of the quantity $1 - f(t)$ when $t \to 1$. Now

$$1 - f(t) = 1 - R(R(t)) = G(R(t)) \sim p_K R(t)^K. \tag{4.30}$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

By definition, $R(x) = 1 - \sum_{k=1}^{\infty} p_k x^k$, thus

$$R'(x) = -\sum_{k=1}^{\infty} k p_k x^{k-1},$$

and again by Theorem 4.7 applied to $a_k = kp_k$ and (1.5),

$$R'(x) \sim -\frac{c\Gamma(1+\rho)}{(1-x)^{\rho}} \text{ as } x \to 1,$$

and thus

$$R(t) = R(1) - \int_t^1 R'(x)\,\mathrm{d}x \sim \frac{c\Gamma(1+\rho)}{1-\rho}(1-t)^{1-\rho}. \tag{4.31}$$

Substituting (4.31) into (4.30), we obtain that

$$1 - f(t) \sim p_K \left(\frac{c\Gamma(1+\rho)}{1-\rho}\right)^K (1-t)^{K(1-\rho)} \text{ as } t \to 1.$$

Thus again, for $f'(1) = \infty$ to hold it is necessary that $K(1-\rho) < 1$.

We have shown that $f(t) \sim C_0 t^{K(1-\rho)}$ as $t \to 0$ and $1 - f(t) \sim C_1(1-t)^{K(1-\rho)}$ as $t \to 1$ for some positive constants C_0, C_1 that we determined explicitly. Note that

$$C_1 = C_0^k p_K^{1-K(1-\rho)}$$

and since $K(1-\rho) < 1$, at least one of the constants C_0, C_1 is different from 1. The proof of Proposition 1.3(c) is completed by the following two lemmas applied as follows: Lemma 4.8 applied with $\alpha = K(1-\rho)$ proves the existence of the distributional limit at either point q with $C_q \neq 1$, and Lemma 4.9 shows that the limit exists at $q = 0$ if and only if it exists at $q = 1$.

Lemma 4.8. *Consider the recursion* (1.3).

(1) *Assume that* $f(t) \sim Ct^{\alpha}$ *with* $C \neq 1$ *and* $\alpha \in (0,1)$ *as* $t \to 0$. *Let* $q_+ = \inf\{x : x > 0,\ x = f(x)\}$. *Then*

$$\mathcal{L}(\alpha^n \log W_{2n} \mid W_{2n} \in [0, q_+]) \xrightarrow{d} V_0, \tag{4.32}$$

where V_0 *is a random variable with* $\mathbb{P}(V_0 \in (-\infty, 0)) = 1$.

(2) *Assume that* $1 - f(t) \sim C(1-t)^{\alpha}$ *with* $C \neq 1$ *and* $\alpha \in (0,1)$ *as* $t \to 1$. *Let* $q_- = \sup\{x : x < 1,\ x = f(x)\}$. *Then*

$$\mathcal{L}(\alpha^n \log(1 - W_{2n}) \mid W_{2n} \in [q_-, 1]) \xrightarrow{d} V_1, \tag{4.33}$$

where V_1 *is a random variable with* $\mathbb{P}(V_1 \in (-\infty, 0)) = 1$.

Lemma 4.9. *Consider the recursion* (1.2). *For* $\alpha \in (0,1)$ *convergence* (4.32) *holds for some* V_0 *with* $\mathbb{P}(V_0 \in (-\infty, 0)) = 1$ *if and only if convergence* (4.33) *holds for some* V_1 *with* $\mathbb{P}(V_1 \in (-\infty, 0)) = 1$.

Note that in Lemma 4.9 we do not assume anything about f; in particular, we do not assume that $C \neq 1$.

Proof of Lemma 4.8. Firstly we show how the second part can be obtained from the first part and then we prove the first part of Lemma 4.8, which corresponds to $q = 0$.

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

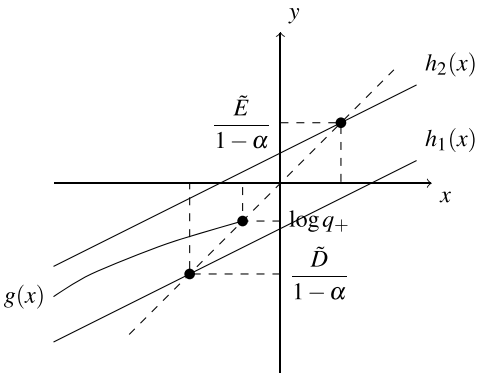


Figure 4. $h_1(x)$, $g(x)$, $h_2(x)$ together with their (stable) fixed points $\tilde{D}/(1-\alpha), \log q_+, \tilde{E}/(1-\alpha)$ respectively. For $x \leqslant \log q_+$, $h_1(x) \leqslant g(x) \leqslant h_2(x)$. The dashed line represents the identity function.

Assume that $1 - f(t) \sim C(1-t)^\alpha$ and set $\tilde{W}_{2n} = 1 - W_{2n}$ and $\tilde{f}(t) = 1 - f(1-t)$. Then,

$$\begin{aligned}\mathbb{P}(\tilde{W}_{2n} \leqslant x) &= \mathbb{P}(1 - W_{2n} \leqslant x)\\ &= 1 - \mathbb{P}(W_{2n} \leqslant 1 - x)\\ &= 1 - f(\mathbb{P}(W_{2n-2} \leqslant 1 - x))\\ &= 1 - f(1 - \mathbb{P}(\tilde{W}_{2n-2} \leqslant x))\\ &= \tilde{f}(\mathbb{P}(\tilde{W}_{2n-2} \leqslant x)) = \cdots = \tilde{f}^n(\mathbb{P}(\tilde{W}_0 \leqslant x))\\ &= \tilde{f}^n(x)\end{aligned}$$

and $\tilde{f}(t) = 1 - f(1-t) \sim Ct^\alpha$ as $t \to 0$. Hence it is enough to prove the result for the case $q = 0$.

Fix some $x < 0$. Note that for n large enough $\exp(x/\alpha^n) \leqslant q_+$. Hence for these n,

$$\begin{aligned}\mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+]) &= \mathbb{P}\left(W_{2n} \leqslant \exp\left(\frac{x}{\alpha^n}\right) \,\middle|\, W_{2n} \in [0, q_+]\right)\\ &= \frac{1}{q_+} f^n\left(\exp\left(\frac{x}{\alpha^n}\right)\right). \qquad (4.34)\end{aligned}$$

Define

$$g(y) = \log(f(\exp(y))),$$

and observe that

$$g^n(y) = \log(f^n(\exp(y))). \qquad (4.35)$$

The idea behind $g(y)$ is to linearize $f(y)$: note that g is a monotone function, $g(\log q_+) = \log q_+$, and that

$$g(y) = \alpha y + O(1) \text{ as } y \to -\infty,$$

hence there exist constants $\tilde{D}, \tilde{E}$ such that, for $y \leqslant \log q_+ < 0$,

$$\tilde{D} + \alpha y \leqslant g(y) \leqslant \tilde{E} + \alpha y. \qquad (4.36)$$

In the first part of the proof we show (assuming that the limit (4.32) exists) that $\mathbb{P}(V \in (-\infty, 0)) = 1$. Define (see Figure 4)

$$\begin{aligned}h_1(y) &= \tilde{D} + \alpha y,\\ h_2(y) &= \tilde{E} + \alpha y.\end{aligned}$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Now

$$\begin{aligned}\lim_{n\to\infty} h_1^n\left(\frac{y}{\alpha^n}\right) &= y + \frac{\tilde{D}}{1-\alpha},\\ \lim_{n\to\infty} h_2^n\left(\frac{y}{\alpha^n}\right) &= y + \frac{\tilde{E}}{1-\alpha},\end{aligned} \tag{4.37}$$

where $\tilde{D}/(1-\alpha)$, $\tilde{E}/(1-\alpha)$ are the (unique) fixed points of h_1 and h_2 respectively.

Equations (4.34), (4.35), (4.36) and (4.37) together imply that

$$\limsup_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leqslant x | W_{2n} \in [0, q_+]) \leqslant \frac{1}{q_+} \exp\left(x + \frac{\tilde{E}}{1-\alpha}\right),$$

and therefore

$$\lim_{x\to-\infty} \limsup_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+]) = 0.$$

We shall now show that

$$\lim_{x\to 0^-} \liminf_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+]) = 1. \tag{4.38}$$

To do so, note first that by (4.34) and (4.35)

$$\mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+]) = \frac{1}{q_+} \exp\left(g^n\left(\frac{x}{\alpha^n}\right)\right),$$

hence (4.38) is equivalent to

$$\lim_{x\to 0^-} \liminf_{n\to\infty} g^n\left(\frac{x}{\alpha^n}\right) = \log q_+. \tag{4.39}$$

Note that $\log q_+$ is a fixed point of $g(x)$. Equation (4.39) indicates that the scaling α^n is not strong enough to compensate the attraction of the fixed point $\log q_+$ of g.

Let $k_{x,n}$ be the smallest k such that

$$h_1^k\left(\frac{x}{\alpha^n}\right) \geqslant \frac{\tilde{D}}{1-\alpha} - 1.$$

Note that $k_{x,n}$ is properly defined as $\tilde{D}/(1-\alpha)$ is the only fixed point of h_1 and is stable. Moreover, by (4.37) we have that, for $x \in (-1, 0)$,

$$\lim_{n\to\infty} h_1^n\left(\frac{x}{\alpha^n}\right) = x + \frac{\tilde{D}}{1-\alpha} \geqslant \frac{\tilde{D}}{1-\alpha} - 1,$$

hence for these x, $n - k_{x,n} \geqslant 0$ for large n. Define also

$$K_x = \liminf_{n\to\infty} (n - k_{x,n})$$

and note that since

$$\lim_{x\to 0^-} \lim_{n\to\infty} h_1^n\left(\frac{x}{\alpha^n}\right) = \frac{\tilde{D}}{1-\alpha},$$

and the right-hand side is a fixed point of h_1, we obtain that

$$\lim_{x\to 0^-} K_x = \infty.$$

Now, similarly, define $\tilde{k}_{x,n}$ to be the smallest k such that

$$g^k\left(\frac{x}{\alpha^n}\right) > \frac{\tilde{D}}{1-\alpha} - 1.$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Since $g(y) \geqslant h_1(y)$ for $y \leqslant \log q_+$, we have $k_{x,n} \geqslant \tilde{k}_{x,n}$, and therefore

$$\lim_{x\to 0^-} \liminf_{n\to\infty} (n - \tilde{k}_{x,n}) = \infty.$$

This implies that

$$\begin{aligned}\lim_{x\to 0^-} \liminf_{n\to\infty} g^n\left(\frac{x}{\alpha^n}\right) &= \lim_{x\to 0^-} \liminf_{n\to\infty} g^{n-k_{x,n}}\left(g^{k_{x,n}}\left(\frac{x}{\alpha^n}\right)\right)\\ &\geqslant \lim_{x\to 0^-} \liminf_{n\to\infty} g^{n-k_{x,n}}\left(\frac{\tilde{D}}{1-\alpha} - 1\right)\\ &= \log q_+.\end{aligned}$$

To finish the proof it is now enough to justify that the limit

$$\lim_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+])$$

exists for all x; recalling (4.34) it is enough to check that the sequence $f^n(\exp(x/\alpha^n))$ is monotone for large n. Since f is a strictly monotone function, the statements

$$f^{n+1}\left(\exp\left(\frac{x}{\alpha^{n+1}}\right)\right) \geqslant f^n\left(\exp\left(\frac{x}{\alpha^n}\right)\right)$$

and

$$f\left(\exp\left(\frac{x}{\alpha^{n+1}}\right)\right) \geqslant \exp\left(\frac{x}{\alpha^n}\right) \tag{4.40}$$

are equivalent. We set $y = x/\alpha^{n+1}$ and $z = \exp(y)$ (therefore $y \to -\infty$ corresponds to $z \to 0$) obtaining that (4.40) is equivalent to

$$f(z) \geqslant z^\alpha.$$

Therefore, if $f(z) \sim Cz^\alpha$ for $C \neq 1$ we note that for each $x < \log q_+$ the sequence $f^n(\exp(x/(\alpha)^n))$ is monotone for n large enough, which yields existence of the limit. This completes the proof of Lemma 4.8. □

Up to now we have only defined W_m for even m. Before we prove Lemma 4.9 we extend it to odd m. Define the distribution of a random variable W_{2n-1} as follows:

$$W_{2n-1} \stackrel{d}{=} \max_{1\leqslant i\leqslant M} W_{2n-2}^{(i)},$$

where M is a random variable drawn from the tree's offspring distribution and $W_{2n-2}^{(i)}$ are independent copies of W_{2n-2} (independent of M). The quantity W_{2n-1} corresponds to the value at the root of a tree of height $2n-1$, with levels alternating between max and min, starting and ending with a max. Similarly we have

$$W_{2n} \stackrel{d}{=} \min_{1\leqslant i\leqslant M} W_{2n-1}^{(i)}.$$

Lemma 4.10 below provides a useful identity which we are going to apply in the proof of Lemma 4.9.

Lemma 4.10. $W_{2n-1} \stackrel{d}{=} G^{-1}(1 - W_{2n-2})$.

Proof of Lemma 4.10. G is the probability generating function of the offspring distribution of the tree, so $G(t) = \mathbb{P}(\max_{1\leqslant i\leqslant M} U_i \leqslant t)$ where U_i are independent uniform random variables and M follows the offspring distribution (independently of $(U_i, i \geqslant 1)$). Decomposing the minimax tree

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

of height $2n-1$ with maximum at levels 1 and $2n-1$, we see that random variables at level $2n-2$ (*i.e.* one level above the leaves) are distributed as $\max_{1\leqslant i\leqslant M} U_i$. Therefore

$$W_{2n-1} \stackrel{d}{=} W_{2n-2}^{\max,G}, \tag{4.41}$$

where $W_{2n-2}^{\max,G}$ is a random variable corresponding to a max-min tree (*i.e.* with maximum at the even levels and minimum at the odd ones) where at the leaves instead of uniform random variables we put random variables with distribution function G. Noting that if U is a uniform random variable then $G^{-1}(U)$ has distribution function G, we see that

$$W_{2n-2}^{\max,G} \stackrel{d}{=} G^{-1}(W_{2n-2}^{\max}). \tag{4.42}$$

Now, since

$$\max_{1\leqslant i\leqslant M} U_i = 1 - \min_{1\leqslant i\leqslant M} (1-U_i) \stackrel{d}{=} 1 - \min_{1\leqslant i\leqslant M} U_i$$

and

$$\min_{1\leqslant i\leqslant M} U_i = 1 - \max_{1\leqslant i\leqslant M} (1-U_i) \stackrel{d}{=} 1 - \max_{1\leqslant i\leqslant M} U_i,$$

we obtain that

$$W_{2n-2}^{\max} \stackrel{d}{=} 1 - W_{2n-2}. \tag{4.43}$$

Finally, combining (4.41), (4.42) and (4.43) completes the proof. □

We are now ready to prove Lemma 4.9.

Proof of Lemma 4.9. The convergence (4.32) is equivalent to the convergence of

$$\lim_{n\to\infty} \mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+]). \tag{4.44}$$

at all the continuity points of the corresponding limiting distribution function and similarly the convergence (4.33) is equivalent to the convergence of

$$\lim_{n\to\infty} \mathbb{P}(\alpha^n \log (1 - W_{2n}) \leqslant x \mid W_{2n} \in [q_-, 1]). \tag{4.45}$$

at all the continuity points of the corresponding limiting distribution function. Fix $x < 0$. For large n,

$$\begin{aligned}
\mathbb{P}(\alpha^n \log W_{2n} \leqslant x \mid W_{2n} \in [0, q_+]) &= \frac{1}{q_+} \mathbb{P}(\alpha^n \log W_{2n} \leqslant x,\ W_{2n} \in [0, q_+]) \\
&= \frac{1}{q_+} \mathbb{P}(W_{2n} \leqslant \exp(x/\alpha^n),\ W_{2n} \in [0, q_+]) \\
&= \frac{1}{q_+} \mathbb{P}(W_{2n} \leqslant \exp(x/\alpha^n)).
\end{aligned}$$

Now by the branching structure of the tree,

$$\mathbb{P}(W_{2n} \leqslant \exp(x/\alpha^n)) = 1 - G(\mathbb{P}(W_{2n-1} > \exp(x/\alpha^n))).$$

Since G is a continuous function, the convergence (4.44) is equivalent to the convergence

$$\lim_{n\to\infty} \mathbb{P}(W_{2n-1} > \exp(x/\alpha^n)).$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

By Lemma 4.10,

$$\begin{aligned}\mathbb{P}(W_{2n-1} > \exp(x/\alpha^n)) &= \mathbb{P}(G^{-1}(1 - W_{2n-2}) > \exp(x/\alpha^n)) \\ &= \mathbb{P}(1 - W_{2n-2} > G(\exp(x/\alpha^n))) \\ &= \mathbb{P}(\alpha^{n-2}\log(1 - W_{2n-2}) > \alpha^{n-2}\log(G(\exp(x/\alpha^n)))) \\ &= 1 - \mathbb{P}(\alpha^{n-2}\log(1 - W_{2n-2}) \leqslant \alpha^{n-2}\log(G(\exp(x/\alpha^n)))).\end{aligned}$$

Since $G(t) \sim p_K t^K$ as $t \to 0$, we observe that

$$\alpha^{n-2}\log(G(\exp(x/\alpha^n))) = \alpha^{n-2}\log(p_K \exp((xK)/\alpha^n) + o(1) = \frac{xK}{\alpha^2} + o(1).$$

This implies that if the convergence (4.45) holds at some point xK/α^2 which is a continuity point of the limiting distribution function, then the convergence (4.44) holds at x. Similarly, if the convergence (4.44) holds at some point x which is a continuity point of the limiting distribution function, then the convergence (4.45) holds at xK/α^2. Since the set of discontinuity points of any distribution function is at most countable, this completes the proof. □

5. Proof of the endogeny result

To prove Theorem 3.1 we use the idea of *bivariate uniqueness* introduced by Aldous and Bandyopadhyay [1].

Informally, the idea is as follows: suppose we allow *two* values at each node. Each coordinate evolves separately, according to the minimax recursions (and using the same realization of the tree structure). If we put bivariate values at the leaves of the tree, we then get a bivariate value at the root of the tree. Let us consider for the moment the case where the values are discrete (as for the Bernoulli case in Theorem 3.1). If the process is endogenous, and the tree is large, then with high probability the two components at the root agree with each other. On the other hand, if the process is not endogenous, then the probability that they disagree stays bounded away from zero as the size of the tree goes to infinity, and in fact we can obtain a bivariate process on the infinite tree which is two-periodic and non-degenerate (in the sense that the two components are not identically the same).

To formalize this we rewrite some of the ideas around (1.2) in new notation.

Let μ be a distribution on $[0, 1]$. We defined $T(\mu)$ to be the distribution of the left-hand side of (1.2), given that the random variables $W_{2n-2}^{(i,j)}$ on the right-hand side of (1.2) are i.i.d. with distribution μ.

So T is a map from $\mathcal{P}$ to $\mathcal{P}$, where $\mathcal{P}$ is the space of distributions on $[0, 1]$. For Theorem 3.1 we assume that the Bernoulli $(1 - x)$ distribution is a fixed point of T.

Now consider the space $\mathcal{P}^{(2)}$ of distributions on $[0, 1]^2$. Define the map $T^{(2)}$ from $\mathcal{P}^{(2)}$ to itself as follows. As before let M and $M_1, M_2, \ldots$ be i.i.d. draws from the offspring distribution. Let $(X_1^{i,j}, X_2^{i,j})$, for each i, j, be i.i.d. with distribution $\mu^{(2)}$ (and independent of M and $\{M_i\}$). Then let $T^{(2)}(\mu^{(2)})$ be the distribution of (X_1, X_2), where

$$X_1 = \min_{1 \leqslant i \leqslant M} \max_{1 \leqslant j \leqslant M_i} X_1^{(i,j)},$$

$$X_2 = \min_{1 \leqslant i \leqslant M} \max_{1 \leqslant j \leqslant M_i} X_2^{(i,j)}.$$

Note particularly that the recursions for X_1 and X_2 use the *same* realization of M and $\{M_i\}$.

If $\mu \in \mathcal{P}$ then we can define a *diagonal* distribution $\mu^{\nearrow}$ on $\mathcal{P}^{(2)}$ by $\mu^{\nearrow} = \mathrm{dist}(X, X)$ if $\mu = \mathrm{dist}(X)$.

If μ is a fixed point of T, then certainly $\mu^{\nearrow}$ is a fixed point of $T^{(2)}$. The question is whether there can be any fixed point of $T^{(2)}$, whose marginals are equal to μ, and which is *not* of the form

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

of the diagonal distribution $\mu^{\nearrow}$. Mach, Sturm and Swart [9, Theorem 1], refining Aldous and Bandyopadhyay [1, Theorem 11], show that the recursive tree process is endogenous if and only if there are no such non-degenerate bivariate fixed points (*i.e.* if the 'bivariate uniqueness property' holds).

Proof of Theorem 3.1. We apply Theorem 1 of [9] (or indeed Theorem 11 of [1], since the additional technical condition relating to the continuity of the operator $T^{(2)}$ does in fact hold in this setting). To prove our result it is enough to show that the bivariate uniqueness property holds if and only if $f'(x) \leqslant 1$.

Let us write μ for the Bernoulli $(1-x)$ distribution on $\{0,1\}$. We look for a distribution $\mu^{(2)}$ on $\{0,1\}^2$ which is a fixed point of $T^{(2)}$, and whose marginals are both μ, but which is not the diagonal distribution $\mu^{\nearrow}$. Once these marginals are specified, we only need to specify one further parameter, say $b = \mu^{(2)}(1,0)$, since then we can deduce $\mu^{(2)}(1,1) = 1 - x - \mu^{(2)}(1,0) = 1 - x - b$, and similarly $\mu^{(2)}(0,1) = b$ and $\mu^{(2)}(0,0) = x - b$. Note that $b \in [0, \min(x, 1-x)]$.

To show that $\mu^{(2)}$ is a fixed point of $T^{(2)}$, again it suffices to check just one entry of $T^{(2)}(\mu^{(2)})$. To look at this we can consider a random tree with two levels, with bivariate marginals according to $\mu^{(2)}$ at level 2 of the tree; we wish to see distribution $\mu^{(2)}$ again at the root. Then write $\nu^{(2)}$ for the corresponding distribution of the marginals at level 1. Let us write o for the root and ι for a typical level 1 node.

So consider the probability of seeing values $(1,0)$ at the root. For this to happen, all children of the root must have 1 in the first coordinate, but at least one child of the root must have 0 in the second coordinate. That is, all children have values $(1,0)$ or $(1,1)$, but not all of them have values $(1,1)$. We obtain

$$\begin{aligned}\mathbb{P}(\text{values } (1,0) \text{ at } o) &= G(\nu^{(2)}(1,0) + \nu^{(2)}(1,1)) - G(\nu^{(2)}(1,1)) \\ &= R(\nu^{(2)}(1,1)) - R(\nu^{(2)}(1,0) + \nu^{(2)}(1,1)). \end{aligned} \tag{5.1}$$

We examine both terms on the right-hand side. First note that $\nu^{(2)}(1,0) + \nu^{(2)}(1,1)$ is the probability that ι has value 1 in the first coordinate. This is the probability that at least one child of ι has value 1 in the first coordinate, *i.e.* that not all the children of ι have value 0 in the first coordinate. Hence

$$\nu^{(2)}(1,0) + \nu^{(2)}(1,1) = 1 - G(\mu(0,1) + \mu(0,0)) = 1 - G(x) = R(x). \tag{5.2}$$

Similarly, for ι to have values $(1,1)$, we need to exclude the two events that all its children have value 0 in the first coordinate or that all its children have value 0 in the second coordinate. Both of these events have probability $G(x)$, while their intersection, *i.e.* that all children have values $(0,0)$, has probability $G(x-b)$. So applying inclusion–exclusion,

$$\nu^{(2)}(1,1) = 1 - G(x) - G(x) + G(x-b) = 2R(x) - R(x-b). \tag{5.3}$$

Combining (5.1), (5.2) and (5.3), we have that if the probability of values $(1,0)$ at level 2 is $b \in [0, \min(x, 1-x)]$, then the probability of values $(1,0)$ at the root is $h(b) \in [0, \min(x, 1-x)]$, where

$$h(b) := \mathbb{R}(2R(x) - R(x-b)) - R(R(x)). \tag{5.4}$$

For $\mu^{(2)}$ to be a fixed point of $T^{(2)}$, we therefore need $b = h(b)$. Also $\mu^{(2)}$ is diagonal if and only if $b = 0$. So non-endogeny is equivalent to the existence of a fixed point of h in the interval $(0, \min(x, 1-x)]$.

From (5.4) we have $h(0) = 0$, and differentiating with respect to b we get

$$h'(b) = R'(R(x) - [R(x-b) - R(x)])R'(x-b) \tag{5.5}$$

so that

$$h'(0) = R'(R(x))R'(x) = \frac{\mathrm{d}}{\mathrm{d}x} R(R(x)) = f'(x).$$

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press

Differentiating once more, we obtain

$$h''(b) = R''(2R(x) - R(x-b))R'(x-b)^2 - R'(2R(x) - R(x-b))R''(x-b). \tag{5.6}$$

Since R is positive, decreasing and strictly concave, it follows that (5.5) is positive and (5.6) is negative, hence that h is increasing and strictly concave.

So if $f'(x) \leqslant 1$, giving $h'(0) \leqslant 1$, then $h(u) < u$ for all $u > 0$. In that case the only non-negative fixed point of h is 0, and we must obtain $b = 0$. In that case the distribution $\mu^{(2)}$ must be a diagonal distribution, and we have bivariate uniqueness (and hence endogeny).

On the other hand, suppose that $f'(x) > 1$, so that $h'(0) > 1$. Then for sufficiently small $\varepsilon > 0$, $h(\varepsilon) > \varepsilon$. Starting from some such ε and iterating h repeatedly gives an increasing sequence which is bounded above by $\min(x, 1-x)$. Its limit is a fixed point of h which lies in $(0, \min(x, 1-x)]$. Hence in this case there does exist a non-degenerate bivariate fixed point, and the process is non-endogenous, as required. □

Proof of Corollary 3.2. Since $f'(x) = 1$ everywhere, Theorem 3.1 tells us that all the processes with Bernoulli marginals are endogenous. This implies that for any μ, for the process with marginals μ, the event $\{Y \leqslant y\}$ is measurable with respect to the structure of the tree, for any y, where Y is the value at the root. But then in fact the random variable Y is measurable with respect to the structure of the tree, as required. □

Acknowledgements

We thank Alexander Holroyd and Julien Berestycki for valuable discussions, and Christina Goldschmidt, Michał Przykucki and the referee for many helpful comments.

References

[1] Aldous, D. J. and Bandyopadhyay, A. (2005) A survey of max-type recursive distributional equations. *Ann. Appl. Probab.* **15** 1047–1110.
[2] Ali Khan, T., Devroye, L. and Neininger, R. (2005) A limit law for the root value of minimax trees. *Electron. Comm. Probab.* **10** 273–281.
[3] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987) *Regular Variation*, Vol. 27 of Encyclopedia of Mathematics and its Applications, Cambridge University Press.
[4] Broutin, N., Devroye, L. and Fraiman, N. (2018) Recursive functions on conditional Galton–Watson trees. arXiv:1805.09425
[5] Broutin, N. and Mailler, C. (2018) And/or trees: A local limit point of view. *Random Struct. Alg.* **53** 15–58.
[6] Browne, C. B., Powley, E., Whitehouse, D., Lucas, S. M., Cowling, P. I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S. and Colton, S. (2012) A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games* **4** 1–43.
[7] Gelly, S., Kocsis, L., Schoenauer, M., Sebag, M., Silver, D., Szepesvári, C. and Teytaud, O. (2012) The grand challenge of computer Go: Monte Carlo tree search and extensions. *Comm. Assoc. Comput. Mach.* **55** 106–113.
[8] Holroyd, A. E. and Martin, J. B. (2019) Galton–Watson games. arXiv:1904.04150
[9] Mach, T., Sturm, A. and Swart, J. M. (2018) A new characterization of endogeny. *Math. Phys. Anal. Geom.* **21** 30.
[10] Pearl, J. (1980) Asymptotic properties of minimax trees and game-searching procedures. *Artif. Intell.* **14** 113–138.
[11] Pemantle, R. and Ward, M. D. (2006) Exploring the average values of Boolean functions via asymptotics and experimentation. In *Proceedings of the Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, SIAM, pp. 253–262.
[12] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. *et al.* (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* **529** (7587) 484–489.

Cite this article: Martin JB and Stasiński R (2020). Minimax functions on Galton-Watson trees. *Combinatorics, Probability and Computing* **29**, 455–484. https://doi.org/10.1017/S0963548319000403

https://doi.org/10.1017/S0963548319000403 Published online by Cambridge University Press