THE NOISY VETO-VOTER MODEL: A RECURSIVE DISTRIBUTIONAL EQUATION ON [0,1]

SAUL JACKA AND MARCUS SHEEHAN

ABSTRACT. We study a particular example of a recursive distributional equation (RDE) on the unit interval. We identify all invariant distributions, the corresponding "basins of attraction" and address the issue of endogeny for the associated tree-indexed problem, making use of an extension of a recent result of Warren.

April 28, 2007

1. INTRODUCTION

Let M be a random variable, taking values in $\overline{\mathbb{N}} = \{1, ...; \infty\}$, and let ξ be an independent Bernoulli(p) random variable.

We consider the following simple Recursive Distributional Equation (henceforth abbreviated as RDE):

$$Y = \xi \prod_{i=1}^{M} Y_i + (1 - \xi)(1 - \prod_{i=1}^{M} Y_i). \tag{1.1}$$

Viewing (1.1) as an RDE, we seek a stationary distribution, ν, such that if Y_i are iid with distribution ν and are independent of (M, ξ), then Y also has distribution ν.

We term (1.1) the noisy veto-voter model since, if each Y_i takes values in $\{0, 1\}$ with value 0 being regarded as a veto, then the outcome is vetoed unless either (a) each voter i 'assents' ($Y_i = 1$ for each $1 \leq i \leq M$) and there's no noise ($\xi = 1$) or (b) someone vetos, but is reversed by the noise ($\xi = 0$).

In this paper, we look for solutions to the RDE (1.1) taking values in $[0, 1]$.

As observed in Aldous and Bandhapadhyay [1], and as we shall explain in a little more detail in section 2, we may think of (families of) solution to the RDE as being located at the nodes of a (family) tree (for a Galton-Watson branching process). Actually, for some purposes we shall find it more convenient to embed this family tree into $\mathbf{T}$, the deterministic tree with infinite branching factor of size $\aleph_0$.

The generic setup in such circumstances is to find distributional fixed points of the recursion:

$$X_u = f(\xi_u; X_{ui}, i \geq 1), \tag{1.2}$$

where X_u and ξ_u are respectively, the value and the noise associated with node u and ui is the address of the ith daughter of node u.

Key words: endogeny; veto-voter model; basin of attraction; Galton-Watson branching process.
AMS 2000 subject classifications: Primary 60E05; secondary 60E99, 60J80.
The authors are grateful for many fruitful discussions with Jon Warren on the topics of this paper.

With this set-up, it is of some interest not only to find solutions to the RDE (1.2) but also to answer the question of endogeny:

'is $(X_u; u \in \mathbf{T})$ measurable with respect to $(\xi_u; u \in \mathbf{T})$?'

If this measurability condition holds, then $X.$ is said to be endogenous.

In this paper we will first show how to transform (1.1) into the new RDE:

$$X = 1 - \prod_{i=1}^{N} X_i, \tag{1.3}$$

for a suitable, random variable N, independent of the X_i, Then we'll not only find all the solutions to this RDE on $[0,1]$, their basins of attractions and the limit cycles of the corresponding map on the space of distributions on $[0,1]$, but also give necessary and sufficient conditions for the corresponding solutions on $\mathbf{T}$ to be endogenous.

2. Notation and a transformation of the RDE

2.1. **Tree-indexed solutions.** We seek distributions ν on $[0,1]$ such that if $(Y_i; 1 \le i)$ are independent with distribution ν, then the random variable Y satisfying (1.1) also has distribution ν. More precisely, writing $\mathcal{P}$ for the set of probability measures on $[0,1]$, suppose that M has distribution d on $\overline{\mathbb{Z}}_+$ and define the map

$$\mathcal{T} \equiv \mathcal{T}_d : \mathcal{P} \to \mathcal{P}$$

Then we set $\mathcal{T}(\nu)$ to be the law of the random variable Y given by (1.1), when the Y_i are independent and identically distributed with distribution ν and are independent of N, and seek fixed points of the map $\mathcal{T}$. The existence and uniqueness of fixed points of this type of map, together with properties of the solutions, are addressed by Aldous and Bandhapadhyay in [1] (the reader is also referred to [2] and [4] and the references therein). The linear and min cases are particularly well-surveyed, though we are dealing with a non-linear case to which the main results do not apply.

A convenient generalisation of the problem is the so-called *tree-indexed* problem, in which we think of the Y_i as being marks associated with the daughter nodes of the root of T, a family tree of a Galton-Watson branching process. We start at some level m of the random tree. Each vertex v in level $m-1$ of the tree has M_v daughter vertices, where the M_v are i.i.d. with common distribution d and has associated with it noise ξ_v, where the $(\xi_u; u \in T)$ are iid and are independent of the $(M_u; u \in T)$.

By associating with daughter vertices independent random variables Y_{vi} having distribution ν, we see that Y_v and $Y_{vi}; 1 \le i \le M_v$ satisfy equation (1.1).

In this setting the notion of endogeny was introduced in [1]. Loosely speaking, a solution to the tree-indexed problem (which we will define precisely in the next section) is said to be endogenous if it is a function of the initial data or noise alone so that no additional randomness is present.

It is convenient to work on a tree with infinite branching factor and then think of the random tree of the previous paragraph as being embedded within it. An initial ancestor (in level zero), which we denote $\emptyset$, gives rise to a countably infinite number of daughter vertices (which form the members of the first generation), each of which gives rise to an infinite number of daughters (which form the members of the second generation), and so on. We assign each vertex an address according to its position in the tree: the members of the first generation are denoted $1, 2, ...$, the second $11, 12, \ldots 21, 22, \ldots, 31, 32, \ldots$ etc, so that vertices in level n of the tree correspond to sequences of positive integers of length n. We also write $uj, j = 1, 2, ...$ for the daughters of a vertex u. We write $\mathbf{T}$ for the collection of all vertices or nodes (i.e. $\mathbf{T} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$) and think of it as being partitioned by depth, that is, as being composed of levels or generations, in the way described and define the depth function $|\cdot|$ by $|u| = n$ if vertex u is in level n of the tree. Associated to each of the vertices $u \in \mathbf{T}$ are iid random variables M_u with distribution d, telling us the (random) number of offspring produced by u. The vertices $u1, u2, ..., uM_u$ are thought of as being alive (relative to $\emptyset$) and the $\{uj : j > M_u\}$ as dead. We can now write our original equation as a recursion on the vertices of $\mathbf{T}$:

$$Y_u = \xi_u \prod_{i=1}^{M_u} Y_{ui} + (1 - \xi_u)(1 - \prod_{i=1}^{M_u} Y_{ui}),\ u \in \mathbf{T}. \tag{2.1}$$

The advantage of the embedding now becomes clear: we can talk about the RDE at any vertex in the infinite tree and yet, because the product only runs over the live daughters relative to u, the random Galton-Watson family tree is encoded into the RDE as noise.

2.2. **The transformed problem.** It is a relatively simple matter to transform the RDE (2.1) into the following, simpler, RDE:

$$X_u = 1 - \prod_{i=1}^{N_u} X_{ui},\ u \in \mathbf{T}. \tag{2.2}$$

To do so, first note that if we colour red all the nodes, v, in the tree $\mathbf{T}$ for which $\xi_v = 0$ then it is clear that we may proceed down each line of descent from a node u until we hit a red node. In this way, we either "cut" the tree at a collection of nodes which we shall view as the revised family of u, or not, in which case u has an infinite family. Denote this new random family size by N_u then

$$Y_u = 1 - \prod_{i=1}^{N_u} Y_{\hat{ui}},$$

if u is red, where $\hat{ui}$ denotes the ith red node in the revised family of u. Now condition on node u being red, then with this revised tree we obtain the RDE (2.2). It is easy to see that if the original tree has family size PGF G, then the family size in the new tree corresponds to the total number of deaths in the original tree when it is independently thinned, with the descendants of each node being pruned with probability q. It is easy to obtain the equation for the PGF, H, of the family size N_u on the new tree:

$$H(z) = G(pH(z) + qz). \tag{2.3}$$

3. The discrete and conditional probability solutions

We begin with some notation and terminology. We say that the random variables in (2.1) are weakly stationary if X_u has the same distribution for every $u \in \mathbf{T}$. The stationarity of the X_u corresponds to X_u having as distribution an invariant measure for the distributional equation (2.2).

Definition 3.1. *We say that the process (or collection of random variables)* $\boldsymbol{X} = (X_u; u \in \boldsymbol{T})$ *is a* tree-indexed *solution to the RDE (2.2) if*

(1) *for every* n*, the random variables* $(X_u; |u| = n)$ *are mutually independent and independent of* $(N_v; |v| \leq n-1)$*;*
(2) *for every* $u \in \boldsymbol{T}$*,* X_u *satisfies*
$$X_u = 1 - \prod_{i=1}^{N_u} X_{ui},$$
and the $(X_u; u \in \boldsymbol{T})$ *are weakly stationary.*

Notice that these conditions determine the law of $\mathbf{X}$. This means that a tree-indexed solution is also stationary in the strong sense, that is, a tree-indexed solution is "translation invariant" with respect to the root (if we consider the collection $\mathbf{X}^v = (X_u; u \in \mathbf{T}_v)$, where $\mathbf{T}_v$ is the sub-tree rooted at v, then $\mathbf{X}^v$ has the same distribution as $\mathbf{X}$ for any $v \in \mathbf{T}$). Furthermore, we say that such a solution is *endogenous* if it is measurable with respect to the random tree (i.e. the collection of family sizes) $(N_u; u \in \mathbf{T})$. As we remarked in the introduction, in informal terms this means that the solution depends only on the noise with no additional randomness coming from the boundary of the tree. See [1] for a thorough discussion of endogeny together with examples.

The following is easy to prove.

Lemma 3.2. *Let* $(X_u; u \in \boldsymbol{T})$ *be a tree-indexed solution to the RDE (2.2). Then the following are equivalent:*

(1) X *is endogenous;*
(2) $X_\emptyset$ *is measurable with respect to* $\sigma(N_u; u \in \boldsymbol{T})$*;*
(3) X_u *is measurable with respect to* $\sigma(N_v; v \in \boldsymbol{T})$ *for each* $u \in \boldsymbol{T}$*;*
(4) X_u *is measurable with respect to* $\sigma(N_v; v \in \boldsymbol{T}_u)$ *for each* $u \in \boldsymbol{T}$*.*

Remark 3.3. *Notice that if a tree-indexed solution to (2.2) is endogenous then property (1) of a tree-indexed solution is automatic: for every* $u \in \boldsymbol{T}$*,* X_u *is measurable with respect to* $\sigma(N_v; v \in \boldsymbol{T}_u)$ *and hence is independent of* $(N_v; |v| \leq n-1)$*.*

Lemma 3.4. *There exists a unique probability measure on* $\{0, 1\}$ *which is invariant under (1.3).*

Proof. Let X be a random variable whose distribution is concentrated on $\{0, 1\}$ and which is invariant under (1.3). Let $\mu^1 = \mathbb{P}(X = 1)$. We have then $\mathbb{P}(X = 0) = 1 - \mu^1$ and
$$\mathbb{P}(X_i = 1;\ \text{for } i = 1, ..., N) = \sum_n \mathbb{P}(X_i = 1;\ \text{for } i = 1, ..., n | N = n)\mathbb{P}(N = n) = H(\mu^1).$$
Now, $X = 0$ if and only if $X_i = 1$ for $i = 1, ..., N$. Hence a necessary and sufficient condition for invariance is
$$1 - \mu^1 = H(\mu^1). \tag{3.1}$$
Now let
$$K(x) \stackrel{def}{=} H(x) + x - 1.$$

Since H is a generating function and $H(0) = 0$, we have $K(0) = -1 < 0$ and $K(1) > 0$ so that K is guaranteed to have a zero in $(0, 1)$, and it is unique since the mapping $x \mapsto H(x) + x$ is strictly increasing. □

We can now deduce that there exists a tree-indexed solution on $\{0,1\}^{\mathbf{T}}$ to the RDE (2.2) by virtue of Lemma 6 of [1].

Theorem 3.5. *Let* $\S = (S_u; u \in \boldsymbol{T})$ *be a tree-indexed solution on* $\{0,1\}^{\boldsymbol{T}}$ *to the RDE (2.2) (i.e. the* S_u *have the invariant distribution on the two point set* $\{0,1\}$*), which we will henceforth refer to as the* discrete solution*. Let* $C_u = \mathbb{P}(S_u = 1 | N_v; v \in \boldsymbol{T})$*. Then* $C = (C_u; u \in \boldsymbol{T})$ *is the unique* endogenous *tree-indexed solution to the RDE.*

Proof. To verify the relationship between the random variables, we have, writing $\mathbf{N} = (N_u; u \in \mathbf{T})$ and $\mathbf{N}_u = (N_v; v \in \mathbf{T}_u)$,

$$\begin{aligned} C_u = \mathbb{P}(S_u = 1|\mathbf{N}) &= \mathbb{E}[1_{(S_u=1)}|\mathbf{N}] = \mathbb{E}[S_u|\mathbf{N}] \\ &= \mathbb{E}[1 - \prod_{i=1}^{N_u} S_{ui}|\mathbf{N}] \\ &= 1 - \mathbb{E}[\prod_{i=1}^{N_u} S_{ui}|\mathbf{N}] \\ &= 1 - \prod_{i=1}^{N_u} \mathbb{E}[S_{ui}|\mathbf{N}] \\ &= 1 - \prod_{i=1}^{N_u} C_{ui}, \end{aligned}$$

since the S_{ui} are independent and $\mathbf{N}$ is strongly stationary. To verify stationarity, let

$$C_u^n = \mathbb{P}(S_u = 1|N_v; |v| \leq n).$$

Then the sequence $(C_u^n)_{n\geq 1}$ is a uniformly bounded martingale and so converges almost surely and in L^2 to a limit which must in fact be C_u. Now, we can write C_u^n as

$$\begin{aligned} C_u^n &= 1 - \prod_{i_1=1}^{N_u} C_{ui_1}^n \\ &= 1 - \prod_{i_1=1}^{N_u} \left(1 - \prod_{i_2=1}^{N_{ui_1}} (...(1 - \prod_{i_{n-1-|u|}=1}^{N_{ui_1 i_2 \dots i_{n-2-|u|}}} (1 - (\mu^1)^{N_{ui_1 i_2 \dots i_{n-1-|u|}}}))...) \right) \\ &\to C_u \text{ a.s.}. \end{aligned} \tag{3.2}$$

This corresponds to starting the distributional recursion at level n of the tree with unit masses at μ^1. Now, $(C_u^n; u \in \mathbf{T})$ is stationary since each C_u^n is the same function of $\mathbf{N}_u$, which are themselves stationary. Since C_u is the (almost sure) limit of a sequence of stationary random variables, it follows that $\mathbf{C} = (C_u; u \in \mathbf{T})$ is stationary. Notice that the conditional probability solution, $\mathbf{C}$, is automatically endogenous since C_u is

$\sigma(N_v; v \in \mathbf{T}_u)$-measurable for every $u \in \mathbf{T}$ and hence $(C_u; |u| = n)$ is independent of $(N_u; |u| \le n-1)$. The independence of the collection $(C_u; |u| = n)$ follows from the fact that the $((S_u, \mathbf{N}_u); |u| = n)$ are independent .

Finally, notice that if $(L_u; u \in \mathbf{T})$ solve the RDE (2.2) and are integrable then $m \stackrel{def}{=} \mathbb{E}L_u$ must satisfy (3.1) and hence must equal μ^1. It now follows, that $L_u^n \stackrel{def}{=} \mathbb{E}[L_u|N_v; |v| \le n] = C_u^n$, since at depth n, $L_u^n = \mu^1$ so that L_u^n also satisfies equation (3.2) and hence must equal C_u^n. Now $L_u^n \to L_u$ a.s. and so, if L is endogenous then it must equal C. This establishes that C is the unique endogenous solution. □

Remark 3.6. *Notice that if $\boldsymbol{S}$ is endogenous then $\boldsymbol{C} = \boldsymbol{S}$ almost surely so that if $\boldsymbol{S}$ and $\boldsymbol{C}$ do not coincide then $\boldsymbol{S}$ cannot be endogenous.*

4. The moment equation and uniqueness of solutions

Many of the results proved in this paper rely heavily on the analysis of equation (4.1) below.

Theorem 4.1. *Any invariant distribution for the RDE (2.2) must have moments $(m_n)_{n\ge0}$ satisfying the equation*

$$H(m_n) - (-1)^n m_n = \sum_{k=0}^{n-1} \binom{n}{k} (-1)^k m_k, \tag{4.1}$$

where $m_n^{1+1/n} \le m_{n+1} \le m_n$ and $m_0 = 1$.

Proof. Let X be a random variable whose distribution is invariant for the RDE and write $m_k = \mathbb{E}[X^k]$. Applying the RDE (2.2) to $(1-X)^n$ we have

$$\mathbb{E}[(1-X)^n] = \mathbb{E}[\prod_{i=1}^{N} X_i^n] = H(m_n).$$

On the other hand, by expanding $(1-X)^n$ we obtain

$$\mathbb{E}[(1-X)^n] = \mathbb{E}[\sum_{k=0}^{n} \binom{n}{k} (-1)^k X_i^k]$$

$$= \sum_{k=0}^{n} \binom{n}{k} (-1)^k m_k,$$

so that

$$H(m_n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k m_k.$$

The condition $m_{n+1} \le m_n$ follows from the fact that the distribution is on $[0,1]$. The other condition follows from the monotonicity of L^p norms. □

As an example, if the random variable N has generating function $H(x) = x^2$ (i.e. $N \equiv 2$), the moment equation tells us that

$$m_1^2 + m_1 - 1 = 0$$

so that $m_1 = (\sqrt{5}-1)/2$. For m_2 we have

$$m_2^2 - m_2 - (2-\sqrt{5}) = 0$$

so that $m_2 = m_1$ or m_1^2 and so on. In fact the two possible moment sequences turn out to be $m_0 = 1, m_n = (\sqrt{5}-1)/2$ for $n \geq 1$ or $m_0 = 1, m_1 = (\sqrt{5}-1)/2, m_n = m_1^n$ for $n \geq 2$.
We suppose from now on that $H(0) = 0$ and H is strictly convex (so that $\mathbb{P}(2 \leq N < \infty) > 0$).
We now state the main result of the paper.

Theorem 4.2. *Let $S = (S_u; u \in \boldsymbol{T})$ and $C = (C_u; u \in \boldsymbol{T})$ be, respectively, the discrete solution and corresponding conditional probability solution to the RDE (2.2). Let $\mu^1 = \mathbb{E}[S_u]$. Then*

1. *S is endogenous if and only if $H'(\mu^1) \leq 1$;*
2. *C is the unique endogenous solution;*
3. *the only invariant distributions for the RDE (2.2) are those of $S_\emptyset$ and $C_\emptyset$.*

The proof of the theorem relies on several lemmas. For (1) we extend a result of Warren [5] by first truncating N and then take limits.

First however, we give some consequences of the moment equation (4.1):

Lemma 4.3. *There are at most two moment sequences satisfying (4.1). Moreover, the first moment m^1 is unique and equal to μ^1, $1 > m^1 > \frac{1}{2}$ and in the case that $H'(m^1) \leq 1$ there is only one moment sequence satisfying (4.1).*

Proof. Uniqueness of μ^1 (the root of $f(m^1) = 1$, where $f : t \mapsto H(t) + t$) has already been shown in Lemma 3.4. Now set

$$g(x) = H(x) - x,$$

then g is strictly convex on [0,1] with $g(0) = 0$ and $g(1-) = H(1-) - 1 \leq 0$. Thus there are at most two solutions of $g(x) = 1 - 2m^1$. Since m^1 itself is a solution, it follows that $1 - 2m^1 \leq 0$ and there is at most one other solution. There is another solution with $m^2 < m^1$ if and only if m^1 is greater than μ^*, the argmin of g, and this is clearly true if and only if $g'(m^1) > 0 \Leftrightarrow H'(m^1) > 1$.

Suppose that this last inequality holds, so that there is a solution, m^2, of $g(x) = 1-2m^1$ with $m^2 < \mu^* < m^1$. There is at most one solution of

$$f(x) = 1 - 3m^1 + 3m^2,$$

and if it exists take this as m^3. Similarly, there is at most one solution of $g(x) = 1 - 4m^1 + 6m^2 - 4m^3$ to the left of μ^* and this is the only possibility for m^4. Iterating the argument, we obtain at most one strictly decreasing sequence $m^1, \ldots$.

□

4.1. **The case of a bounded branching factor.** Recall that the random family size N may take the value ∞.

Lemma 4.4. *Define $N^n = min(n, N)$ and denote its generating function by H_n. Then N^n is bounded and*

(1) *$H_n(s) \geq H(s)$ for all $s \in [0,1]$;*
(2) *$H_n \to H$ uniformly on compact subsets of $[0,1)$;*
(3) *$H'_n \to H'$ uniformly on compact subsets of $[0,1)$.*

We leave the proof to the reader.

The following lemma will be used in the proof of Theorem 4.7.

Lemma 4.5. *Let $C_u^{(n)} = \mathbb{P}(S_u = 1 | N_u^n; u \in \boldsymbol{T})$ denote the conditional probability solution for the RDE (2.2) with N replaced by N^n. Let $\mu_n^k = \mathbb{E}[(C_u^{(n)})^k]$ denote the corresponding kth moment and let $\mu^k = \mathbb{E}[(C_u)^k]$. Let μ_n^* denote the argmin of $g_n(x) \stackrel{def}{=} H_n(x) - x$ and let $\mu_{n,m}^2$ denote that root of the equation,*

$$g_n(x) = 1 - \mu_m^1 - \mu_n^1, \tag{4.2}$$

which lies to the left of μ_n^* (i.e. the lesser of the two possible roots). Then $\mu_n^k \to \mu^k$ for $k = 1, 2$ and $\mu_{n,m}^2 \to \mu^2$ as $\min(n, m) \to \infty$.

Proof. For the case $k = 1$, consider the graphs of the functions $H_n(x) + x$ and $H(x) + x$. We have $H_n(x) \geq H(x)$ for all $x \geq 0$ and for all $n \geq 1$ so that μ_n^1 is bounded above by μ^1 for every n, since μ_n^1 and μ^1 are respectively the roots of

$$H_n(x) + x = 1 \text{ and } H(x) + x = 1.$$

Furthermore, since H_n decreases to H pointwise on $[0,1)$, it follows that the μ_n^1 are increasing. The μ_n^1 must therefore have a limit, which we will denote $\widehat{\mu}$.

It follows from Lemma 4.4 that, since $\mu^1 < 1$, $H_n(\mu_n^1) \to H(\widehat{\mu})$. Hence

$$1 = H_n(\mu_n^1) + \mu_n^1 \to H(\widehat{\mu}) + \widehat{\mu},$$

so that $\widehat{\mu}$ is a root of $H(x) + x = 1$. It follows, by uniqueness, that $\widehat{\mu} = \mu^1$.

For the case $k = 2$ we consider the graphs of $g_n(x)$ and $g(x)$. We first show that $\mu_n^2 \to \mu^2$ and then that $\mu_{n,m}^2 \to \mu^2$ as $\min(n, m) \to \infty$.

To show that $\mu_n^2 \to \mu^2$ we argue that μ^2 is the only limit point of the sequence $(\mu_n^2)_{n \geq 1}$. Notice that, since $\mu_n^1 \to \mu^1$ and μ_n^2 satisfies

$$H_n(\mu_n^2) - \mu_n^2 = 1 - 2\mu_n^1,$$

the only possible limit points of the sequence $(\mu_n^2)_{n \geq 1}$ are μ^1 and μ^2. Now, either $\mu^1 \leq \mu^*$, in which case $\mu^1 = \mu^2$ or, $\mu^2 \leq \mu^* < \mu^1 < 1$. In the latter case, it is easy to show that $\mu_n^* \to \mu^*$ (by uniform continuity of g'_n) and so, since $\mu_n^1 \to \mu^1$ it follows that

$$\mu_n^1 > \mu_n^*,$$

for sufficiently large n, and hence

$$\mu_n^2 \leq \mu_n^*,$$

for sufficiently large n. In either case, the only possible limit point is μ^2; since the μ_n^2 are bounded they must, therefore, converge to μ^2.

We conclude the proof by showing that μ^2 is the only limit point of the sequence $(\mu_{n,m}^2)$.

Since $\mu_m^1, \mu_n^1 \to \mu^1$ as $\min(n,m) \to \infty$ and $\mu_{n,m}^2$ satisfies (4.2), the only possible limit points of the sequence $(\mu_{n,m}^2)_{m,n\geq 1}$ are μ^1 and μ^2.

Once more, consider the two cases:

$$\mu^1 \leq \mu^* \text{ and } \mu^1 > \mu^*.$$

In the first case, $\mu_n^1 = \mu_n^2$, for sufficiently large n, so that μ^2 is the only limit point; in the second case

$$\mu^1 = \liminf_n \mu_n^1 > \mu^* = \limsup_n \mu_n^*,$$

and since $\mu_{n,m}^2 \leq \mu_n^*$, μ^1 cannot be a limit point. Thus, in either case, μ^2 is the unique limit point and hence is the limit.

□

Remark 4.6. *Notice that the method of the proof can be extended to prove that $\mu_n^k \to \mu^k$ for any k.*

Theorem 4.7. *$C_u^{(n)}$ converges to C_u in L^2.*

Proof. Let $n \geq m$. Define $E_{m,n} = \mathbb{E}[(C_u^{(m)} - C_u^{(n)})^2]$. Expanding this, we obtain

$$E_{m,n} = \mu_m^2 + \mu_n^2 - 2r_{m,n},$$

where $r_{m,n} = \mathbb{E}[C_u^{(m)} C_u^{(n)}]$. On the other hand, by applying the RDE (2.2) once, we obtain

$$E_{m,n} = \mathbb{E}[(\prod_{i=1}^{N_u^n} C_{ui}^{(n)} - \prod_{i=1}^{N_u^m} C_{ui}^{(m)})^2]$$

$$= H_m(\mu_m^2) + H_n(\mu_n^2) - 2\mathbb{E}[\prod_{i=1}^{N_u^m} C_{ui}^{(m)} \prod_{i=1}^{N_u^n} C_{ui}^{(n)}].$$

We can bound $E_{m,n}$ above and below as follows: since each C_{ui}^k is in [0,1] omitting terms from the product above increases it, while adding terms decreases it. Thus, since $n \geq m$, $N_u^n \geq N_u^m$, and so replacing N_u^n by N_u^m in the product above increases it while replacing N_u^m by N_u^n decreases it. Thus we get:

$$H_m(\mu_m^2) + H_n(\mu_n^2) - 2H_m(r_{m,n}) \leq E_{m,n} \leq H_m(\mu_m^2) + H_n(\mu_n^2) - 2H_n(r_{m,n}).$$

Using the upper bound we have

$$2H_n(r_{m,n}) \leq H_m(\mu_m^2) + H_n(\mu_n^2) - E_{m,n} = H_m(\mu_m^2) + H_n(\mu_n^2) - \mu_m^2 - \mu_n^2 + 2r_{m,n}.$$

The moment equation (4.1) tells us that $H_m(\mu_m^2) - \mu_m^2 = 1 - 2\mu_m^1$ and that $H_n(\mu_n^2) - \mu_n^2 = 1 - 2\mu_n^1$. Hence

$$2H_n(r_{m,n}) \leq 1 - 2\mu_m^1 + \mu_m^2 + 1 - 2\mu_n^1 + \mu_n^2 - \mu_m^2 - \mu_n^2 + 2r_{m,n},$$

so that, on simplifying,

$$H_n(r_{m,n}) - r_{m,n} \leq 1 - \mu_m^1 - \mu_n^1.$$

Recall that the equation $H_n(x) - x = 1 - \mu_m^1 - \mu_n^1$ has (at most) two roots, the lesser of which we denoted $\mu_{m,n}^2$. Let $\mu_{m,n}^1$ be the other (larger) root (or 1, if the second root does not exist). Then, since $H_n(x) - x$ is convex, $\mu_{n,m}^2 \leq r_{m,n} \leq \mu_{n,m}^1$ for all m, n and hence $\liminf_{m\to\infty} r_{m,n} \geq \mu^2$ since $\mu_{n,m}^2 \to \mu^2$ by Lemma 4.5.

On the other hand, Holder's inequality tells us that $r_{m,n} \le \sqrt{\mu_m^2 \mu_n^2}$ and so it follows that $\limsup_{m\to\infty} r_{m,n} \le \mu^2$ since $\mu_m^2, \mu_n^2 \to \mu^2$ by Lemma 4.5. Hence $r_{m,n} \to \mu^2$ as $n \to \infty$ and

$$E_{m,n} \to \lim_{m,n\to\infty} \mu_m^2 + \mu_n^2 - 2r_{m,n} = \mu^2 + \mu^2 - 2\mu^2 = 0,$$

showing that $(C_u^{(n)})$ is Cauchy in L^2. It now follows, by the completeness of L^2, that $C_u^{(n)}$ converges. Since $C_u^{(n)}$ is $\sigma(N)$-measurable, the limit L_u of the $C_u^{(n)}$ must also be $\sigma(N)$-measurable for each u and the collection $(L_{ui})_{i\ge1}$ must be independent and identically distributed on [0,1], with common mean $\mu^1 < 1$. Moreover, by strong stationarity of the $C^{(n)}$s, the L_us are strongly stationary.

To verify that $L_\emptyset$ is the conditional probability solution, notice that

$$1_{E_n} C_\emptyset^{(n)} = (1 - \prod_{i=1}^{N_\emptyset^n} C_i^{(n)}) 1_{E_n}$$

$$= (1 - \prod_{i=1}^{N_\emptyset} C_i^{(n)}) 1_{E_n},$$

where $E_n = \{N_\emptyset \le n\}$. As $n \to \infty$, $E_n \uparrow E \overset{def}{=} (N < \infty)$; furthermore, since the $C_i^{(n)}$ converge in L^2, they do so in probability. We may assume without loss of generality, therefore, that $C_i^{(n)}$ converges almost surely for each i so that, in the limit,

$$1_E L_\emptyset = \lim 1_{E_n} C_\emptyset^{(n)} = \lim 1_{E_n} (1 - \prod_{i=1}^{N_\emptyset} C_i^{(n)}) = 1_E (1 - \prod_{i=1}^{N_\emptyset} L_i) \text{ a.s.} \tag{4.3}$$

It is easy to show that

$$\prod_{i=1}^{\infty} L_i = 0 \text{ a.s.}$$

while

$$1_{E^c} C_\emptyset^{(n)} = 1_{E^c} (1 - \prod_{i=1}^{(n)} C_i^{(n)}) \to 1_{E^c} \text{ a.s.},$$

so that

$$1_{E^c} L_\emptyset = \lim 1_{E^c} C_\emptyset^{(n)} = 1_{E^c}. \tag{4.4}$$

Thus adding equations (4.3) and (4.4) we see that

$$L_\emptyset = (1 - \prod_{i=1}^{N_\emptyset} L_i),$$

and so L is an endogenous solution to the RDE. It follows from uniqueness that L must be the conditional probability solution C. □

4.2. **Proof of Theorem 4.2.** We are now nearly in a position to finish proving Theorem 4.2. To recap, we have shown in Lemma 4.3 that there are at most two distributions which solve the RDE (1.3), corresponding to the 'moment sequences' $\mu^1, \mu^1, \ldots$ and $\mu^1, \mu^2, \ldots$. The first of these is the moment sequence corresponding to the distribution on $\{0, 1\}$ with mass μ^1 at 1. The second may or may not be a true moment sequence and is equal to the first if and only if $H'(\mu^1) \leq 1$. Moreover, there is only one endogenous solution, and this corresponds to the conditional probability solution C, thus if we can show that C is not discrete (i.e. is not equal to S) whenever $H'(\mu^1) > 1$ then we will have proved the result.

We need to recall some theory from [5]. Consider the recursion

$$\xi_u = \phi(\xi_{u0}, \xi_{u1}, ..., \xi_{u(d-1)}, \epsilon_u), \qquad u \in \Gamma_d,$$

where the ξ_u take values in a finite space $\mathcal{S}$, the "noise" terms ϵ_u take values in a space E, Γ_d is the deterministic d-ary tree and ϕ is symmetric in its first $d-1$ arguments. We suppose that the ϵ_u are independent with common law ν and that there exists a measure π which is invariant for the above recursion (i.e. π is a solution of the associated RDE). Let $u_0 = \emptyset, u_1, u_2, \ldots$ be an infinite sequence of vertices starting at the root, with u_{n+1} being a daughter of u_n for every n. For $n \leq 0$, define $\xi_n = \xi_{u_{-n}}$. Then, under the invariant measure π, the law of the sequence $(\xi_n; n \leq 0)$, which, by the symmetry of ϕ does not depend on the choice of sequence of vertices chosen, is that of a stationary Markov chain. Let P^2 be the transition matrix of a Markov chain on $\mathcal{S}^2$, given by

$$P^2((x_1, x_1'), A \times A') = \int_{\mathcal{S}} \int_E 1_(\phi(x_1, x_2, ..., x_d, z) \in A, \phi(x_1', x_2, ..., x_d, z) \in A') d\nu(z) d\pi(x_2) ... d\pi(x_d).$$

Let P^- be the restriction of P^2 to non-diagonal terms and ρ the Perron- Frobenius eigenvalue of the matrix corresponding to P^-.

The following theorem gives a necessary and sufficient condition for endogeny of the tree-indexed solution corresponding to μ. This is a small generalisation of Theorem 1 of [5]

Theorem 4.8. *The tree-indexed solution to the RDE associated with*

$$\xi_u = \phi(\xi_{u0}, \xi_{u1}, ..., \xi_{u(d-1)}, \epsilon_u),$$

corresponding to the invariant measure π, is endogenous if $d\rho < 1$; it is non-*endogenous if $d\rho > 1$. In the critical case $d\rho = 1$, let $\mathcal{H}_0$ be the collection of L^2 random variables measurable with respect to $\xi_\emptyset$ and let $\mathcal{K}$ denote the L^2 random variables measurable with respect to $(\epsilon_u; u \in \Gamma_d)$. Then endogeny holds in this case provided P^- is irreducible and $\mathcal{H}_0 \cap \mathcal{K}^\perp = \{0\}$. See* [5] *for full details.*

Theorem 4.9. *Consider the RDE*

$$X_u = 1 - \prod_{i=1}^{N_u^n} X_{ui}. \tag{4.5}$$

Then, by Lemma 3.4, there exists an invariant probability measure on $\{0, 1\}$ for (4.5). Let μ_n^1 denote the probability of a 1 under this invariant measure. Then the corresponding tree-indexed solution is endogenous if and only if $H_n'(\mu_n^1) \leq 1$.

Proof. Let $N^* = ess\sup N < \infty$ be a bound for N. We can then think of the random tree with branching factor N as being embedded in a N^*-ary tree. Each vertex has N^* daughter vertices and the first N of these are thought of as being *alive* (the remaining being *dead*). In this context our RDE reads

$$X = 1 - \prod_{live\ u} X_u.$$

We now compute the transition probabilities from the previous theorem. Consider first the transition from $(0,1)$ to $(1,0)$. The first coordinate automatically maps to 1 and the second maps to 0 provided all of the inputs not on the distinguished line of descent are equal to 1. The conditional probability of the vertex on the distinguished line of descent being alive is N/N^* since there are N^* vertices, of which N are alive. The probability of the remaining $N-1$ vertices each taking value 1 is $(\mu_n^1)^{N-1}$ and so the probability of a transition from $(0,1)$ to $(1,0)$, conditional on N, is just

$$1_{(N\geq 1)} \frac{(\mu_n^1)^{N-1} N}{d}.$$

Taking expectations, the required probability is

$$\mathbb{E}[1_{(N\geq 1)} \frac{(\mu_n^1)^{N-1} N}{N^*}] = \frac{\mathbb{E}[1_{(N\geq 1)} N (\mu_n^1)^{N-1}]}{N^*} = \frac{H_n'(\mu_n^1)}{N^*}.$$

The probability of a transition from $(1,0)$ to $(0,1)$ is the same by symmetry. Hence P^- is given by

$$P^- = \begin{pmatrix} 0 & \frac{H_n'(\mu_n^1)}{N^*} \\ \frac{H_n'(\mu_n^1)}{N^*} & 0 \end{pmatrix},$$

and the Perron-Frobenius eigenvalue ρ is $\frac{H_n'(\mu_n^1)}{N^*}$. By Theorem 4.8, the criterion for endogeny is $N^*\rho \leq 1$, i.e. $H_n'(\mu_n^1) \leq 1$, provided that, in the critical case $H_n'(\mu_n^1) = 1$, we verify the stated non-degeneracy conditions.

It is easily seen that P^- is irreducible. For the other criterion, let $X \in \mathcal{H}_0 \cap \mathcal{K}^\perp$ so that $X = f(X_\emptyset)$ for some L^2 function f and $\mathbb{E}[XY] = 0$ for all $Y \in \mathcal{K}$. Taking $Y = 1$, we obtain $\mathbb{E}[X] = 0$. Writing X as

$$X = a1_{(X_\emptyset=1)} + b1_{(X_\emptyset=0)},$$

where a, b are constants, we obtain

$$X = a1_{(X_\emptyset=1)} - \frac{a\mu_n^1}{1-\mu_n^1} 1_{1(X_\emptyset=0)}.$$

For convenience we will scale by taking $a = 1$ (we assume that $X \neq 0$):

$$X = 1_{(X_\emptyset=1)} - \frac{\mu_n^1}{1-\mu_n^1} 1_{(X_\emptyset=0)}.$$

Now, for each k take $Y_k = 1_{(N_\emptyset=k)} \in \mathcal{K}$. Then

$$\mathbb{E}[XY_k] = \mathbb{E}[1_{(N_\emptyset=k)}(1_{(X_\emptyset=1)} - \frac{\mu_n^1}{1-\mu_n^1} 1_{(X_\emptyset=0)})]$$

$$= \mathbb{P}(N = k)[1 - (\mu_n^1)^k - \frac{(\mu_n^1)^{k+1}}{1 - \mu_n^1}]$$

$$= \mathbb{P}(N = k)(1 - \frac{(\mu_n^1)^k}{1 - \mu_n^1}).$$

Now if we sum this expression over k we get $1 - \frac{H_n(\mu_n^1)}{1-\mu_n^1} = 0$. So either each term in the sum is zero or one or more are not. But at least two of the probabilities are non-zero by assumption (at least for sufficiently large n) whilst the term $(1 - \frac{(\mu_n^1)^k}{1-\mu_n^1})$ can only disappear for at most one choice of k. Hence at least one of the terms is non-zero and this contradicts the assumption that $X \in \mathcal{H}_0 \cap \mathcal{K}^\perp$. □

Proof of the remainder of Theorem 4.2 We prove that $H'(\mu^1) > 1$ implies **S** is not endogenous so that **C** cannot equal **S**.

By Theorem 4.9 we know that the RDE (4.5) has two invariant distributions if and only if $H_n'(\mu_n^1) > 1$. But we know that $C_u^{(n)}$ converges to C_u in L^2 and hence $\mu_n^2 \to \mu^2 \neq \mu^1$ so that S_u and C_u have different second moments. It now follows that S_u does not have the same distribution as C_u. Since $[0, 1]$ is bounded, this sequence of moments determines a unique distribution which is therefore that of C_u: see Theorem 1 of Chapter VII.3 of Feller [3] □

5. Basins of attraction

Now we consider the *basin of attraction* of the endogenous solution. That is, we ask for what initial distributions does the corresponding solution at root, $X_\emptyset$, converge (in law) to the endogenous solution.

Definition 5.1. *Let ς be the law of the endogenous solution. Suppose that we insert independent, identically distributed random variables with law ν at level n of the tree and apply the RDE to obtain the corresponding solution $X_u^n(\nu)$ (with law $\mathcal{T}^{n-|u|}(\nu)$) at vertex u.*

The basin of attraction $B(\pi)$ of any solution is given by

$$B(\pi) = \{\nu \in \mathcal{P} : \mathcal{T}^n(\nu) \overset{weak^*}{\to} \pi\},$$

which is, of course, equivalent to the set of distributions ν for which $X_u^n(\nu)$ converges in law to a solution X of the RDE, with law π.

5.1. The unstable case: $H'(\mu^1) > 1$.

Lemma 5.2. *Suppose that $H'(\mu^1) > 1$. Then $X_u^n(\nu) \overset{L^2}{\to} C_u$, the endogenous solution, for any ν with mean μ^1 other than the discrete measure on $\{0, 1\}$.*

Proof. Let $E_k = \mathbb{E}[X_u^n(\nu)^2]$, where $k = n - |u|$, and let $r_k = \mathbb{E}[C_u X_u^n(\nu)]$. Then

$$\mathbb{E}[(X_u^n(\nu) - C_u)^2] = E_k - 2r_k + \mu^2.$$

Now,

$$E_k = \mathbb{E}[(1 - 2\prod_{i=1}^{N_u} X_{ui}^n(\nu) + \prod_{i=1}^{N_u} X_{ui}^n(\nu)^2)]$$

$$= 1 - 2H(\mu^1) + H(E_{k-1}).$$

This is a recursion for E_k with at most two fixed points (recall that the equation $H(x) - x = \mathit{constant}$ has at most two roots). Recalling the moment equation (4.1), these are easily seen to be μ^1 and μ^2, the first and second moments of the endogenous solution. We have assumed that ν is not the discrete distribution and so its second moment (i.e. E_0) must be strictly less than μ^1. Now, under the assumption that $H'(\mu^1) > 1$, μ^1 and μ^2 lie either side of the minimum μ^* of $H(x) - x$ and $H'(\mu^*) = 1$ so that $H'(\mu^2) < 1$. Hence μ^2 is the stable fixed point and it now follows that E_k converges to μ^2.

The recursion for r_k is essentially the same as that for E_k:

$$\mu^2 - r_k = H(\mu^2) - H(r_{k-1}).$$

This has μ^1 and μ^2 as fixed points and, since

$$r_0 = \mathbb{E}[C_u X_u(\nu)] \leq \sqrt{\mathbb{E}[C_u^2]\mathbb{E}[X_u(\nu)^2]} < \sqrt{\mu^1\mu^1} = \mu^1,$$

we are in the same situation as with E_k. That is, we start to the left of μ^1 and, because $H'(\mu^1) > 1$, we conclude that μ^1 is repulsive and it follows that r_k converges to μ^2 under the assumptions of the lemma. Hence

$$\mathbb{E}[(X_u^n(\nu) - C_u)^2] = E_k - 2r_k + \mu^2 \to 0.$$

□

Theorem 5.3. *Let δ denote the discrete distribution on $\{0, 1\}$ with mean μ^1. Then*

$$B(\varsigma) = \{\nu \in \mathcal{P} : \int x d\nu(x) = \mu^1 \textit{ and } \nu \neq \delta\}.$$

That is, $B(\varsigma)$ is precisely the set of distributions on $[0, 1]$ with the correct mean (except the discrete distribution with mean μ^1).

Proof. We have already shown that

$$\{\nu \in \mathcal{P} : \int x d\nu(x) = \mu^1 \textit{ and } \nu \neq \delta)\} \subseteq B(\varsigma).$$

Since the identity is bounded on $[0, 1]$, we conclude that

$$\mathbb{E}X_u^n(\nu) \to \mathbb{E}C_u, \text{ if } \nu \in B(\varsigma),$$

so that $\nu \in B(\varsigma)$ only if the mean of $\mathcal{T}^n(\nu)$ converges to μ^1. From the moment equation (4.1), the mean of $X_u^n(\nu)$ is obtained by iterating the map f n times, starting with the mean of ν. This mapping has a unique fixed point μ^1 and, since $H'(\mu^1) > 1$, it is repulsive. It follows that the only way we can have convergence in mean is if we start with the correct mean, that is, if ν has mean μ^1. Hence

$$B(\varsigma) \subseteq \{\nu \in \mathcal{P} : \int x d\nu(x) = \mu^1 \textit{ and } \nu \neq \delta\}.$$

□

5.2. The stable case: $H'(\mu^1) \leq 1$.

Theorem 5.4. *Let $b(\mu^1)$ be the basin of attraction of μ^1 under the iterative map for the first moment, $f : t \mapsto 1 - H(t)$. Then*

$$B(\varsigma) = \{\nu \in \mathcal{P} : \int x d\nu(x) \in b(\mu^1)\}.$$

Consider once again $\mathbb{E}[(X_u^n(\nu) - C_u)^2]$. Let $m_k^\theta = \mathbb{E}X_u^n(\nu)^\theta$, where $k = n - |u|$. Then

$$m_k^2 = \mathbb{E}(1 - 2\prod_{i=1}^{N_u} X_{ui}^n(\nu) + \prod_{i=1}^{N_u} X_{ui}^n(\nu)^2)$$

$$= 1 - 2H(m_{k-1}^1) + H(m_{k-1}^2).$$

Recalling that $r_k = \mathbb{E}[C_u X_u^n(\nu)]$, we have

$$r_k = \mathbb{E}[(1 - \prod_{i=1}^{N_u} C_{ui})(1 - \prod_{i=1}^{N_u} X_{ui}^n(\nu))]$$

$$= \mathbb{E}[(1 - \prod_{i=1}^{N_u} C_{ui} - \prod_{i=1}^{N_u} X_{ui}^n(\nu) + \prod_{i=1}^{N_u} C_{ui} X_{ui}^n(\nu))]$$

$$= 1 - H(\mu^1) - H(m_{k-1}^1) + H(r_{k-1}).$$

We now turn our attention to analysing the dynamics of m_k^2 and r_k. We will concentrate on the equation for m_k^2 as the equation for r_k is essentially the same. By assumption, m_k^1 converges to μ^1 and so we may approximate m_k^1, for $k \geq k_\epsilon$ (say), by $\mu^1 \pm \epsilon$, for some small $\epsilon > 0$.

Lemma 5.5. *The trajectory l_k of the dynamical system defined by the recursion*

$$l_k = 1 - 2H(\mu^1 + \epsilon) + H(l_{k-1}), \quad l_{k_\epsilon} = m_{k_\epsilon}^2,$$

is a lower bound for m_k^2 for all $k \geq k_\epsilon$, where k_ϵ is a positive integer chosen so that

$$|m_k^1 - \mu^1| < \epsilon, \text{ for } k \geq k_\epsilon.$$

The proof is obvious.

Lemma 5.6. *Let*

$$f_\epsilon(x) = 1 - 2H(\mu^1 + \epsilon) + H(x), \qquad x \in [0, 1].$$

Then, for sufficiently small $\epsilon > 0$, f_ϵ has a unique fixed point $\mu^1(\epsilon)$ for which $\mu^1(\epsilon) < \mu^*$. Moreover, as $\epsilon \to 0$, $\mu^1(\epsilon) \to \mu^1$.

Proof. This follows from uniform continuity, the fact that $H(\mu^1 + \epsilon) < H(\mu^1)$ and the the fact that $H'(\mu^1) \leq 1 \Rightarrow \mu^1 \leq \mu^*$. □

Lemma 5.7. *l_k converges to $\mu^1(\epsilon)$.*

Proof. We have $l_k = f_\epsilon^{k-k_\epsilon}(l_{k_\epsilon})$ and so we need only verify that l_{k_ϵ} is in the basin of attraction of $\mu^1(\epsilon)$ and that $\mu^1(\epsilon)$ is stable. We know that

$$f_\epsilon(\mu^1+\epsilon) < \mu^1+\epsilon$$

since $1-H(\mu^1+\epsilon) < 1-H(\mu^1) = \mu^1$ and so it must be the case that $\mu^1+\epsilon \in (\mu^1(\epsilon), p(\epsilon))$. It now follows that $l_{k_\epsilon} < p(\epsilon)$ since $l_{k_\epsilon} \le m^1_{k_\epsilon} < \mu^1+\epsilon$. In the strictly stable case $H'(\mu^1) < 1$, the stability of $\mu^1(\epsilon)$ follows from the fact that $\mu^1(\epsilon)$ converges to μ^1 as ϵ tends to zero (by the previous lemma) and therefore $\mu^1(\epsilon)$ can be made arbitrarily close to μ^1 by choosing ϵ to be sufficiently small. This means that for sufficiently small ϵ, $H'(\mu^1(\epsilon)) < 1$ by the continuity of H'. In the critical case $H'(\mu^1) = 1$, we have $\mu^1(\epsilon) < \mu^1$, so that (by strict convexity) $H'(\mu^1(\epsilon)) < 1$. In either case it now follows that $f_\epsilon^{k-k_\epsilon}(l_{k_\epsilon})$ converges to $\mu^1(\epsilon)$. □

Proof of Theorem 5.4 The preceding lemmas tell us that

$$\liminf_{k\to\infty} m_k^2 \ge \lim_{k\to\infty} l_k = \mu^1(\epsilon).$$

Letting ϵ tend to zero, we obtain

$$\liminf_{k\to\infty} m_k^2 \ge \mu^1.$$

The fact that $m_k^2 \le m_k^1$ for every k gives us the corresponding inequality for the lim sup:

$$\limsup_{k\to\infty} m_k^2 \le \lim_{k\to\infty} m_k^1 = \mu^1.$$

We conclude that m_k^2 converges to μ^1.

Now,

$$\mathbb{E}[(X_u^n(\nu) - C_u)]^2 = m_k^2 - 2r_k + \mu^2,$$

so that $\mathbb{E}[(X_u^n(\nu) - C_u)^2] \to 0$, remembering that in the stable case the discrete solution and endogenous solution coincide (i.e. $\mu^1 = \mu^2$). We have now shown that

$$\{\nu \in \mathcal{P} : \int x d\nu(x) \in b(\mu^1)\} \subseteq B(\varsigma),$$

and the necessity for convergence in mean ensures that we have the reverse inclusion. This completes the proof. □

6. Outside the basin of attraction of the endogenous solution

In this section we examine what happens if we iterate distributions with mean outside the basin of attraction of the endogenous solution.

Definition 6.1. *Recall that a map f has an n-cycle starting from p if $f^n(p) = p$, where f^n denotes the n-fold composition of f with itself.*

It is easily seen that the map for the first moment $f : t \mapsto 1 - H(t)$ can have only one- and two-cycles. This is because the iterated map $f^{(2)} : t \mapsto 1 - H(1 - H(t))$ is increasing in t and hence can have only one-cycles. Notice also that the fixed points (or one-cycles) of $f^{(2)}$ come in pairs: if p is a fixed point then so too is $1 - H(p)$.

We consider the iterated RDE:

$$X = 1 - \prod_{i=1}^{N_\emptyset}(1 - \prod_{j=1}^{N_i} X_{ij}). \tag{6.1}$$

This corresponds to the iterated map on laws on [0,1], $\mathcal{T}^2$, where $\mathcal{T}$ is given at the beginning of section 2. We denote a generic two-cycle of the map $f^{(2)}$ by the pair (μ^1_+, μ^1_-).

Theorem 6.2. *Suppose that (μ^1_+, μ^1_-) is a two-cycle of $f^{(2)}$. There are at most two solutions of the RDE (6.1) with mean μ^1_+. There is a unique endogenous solution C^+, and a (possibly distinct) discrete solution, S^+, taking values in $\{0,1\}$. The endogenous solution C^+ is given by $P(S^+ = 1|\boldsymbol{T})$ (just as in the non-iterated case). The solutions are distinct if and only if $H'(\mu^1_-)H'(\mu^1_+) > 1$, i.e. if and only if μ^1_+ (or μ^1_-) is an unstable fixed point of $f^{(2)}$.*

Proof. This uses the same method as the proofs of results in section 4.

First, it is clear that S^+ is a solution to (6.1), where $P(S^+ = 1) = \mu^1_+ = 1 - P(S^+ = 0)$. Now take interleaved tree-indexed solutions to the RDE on the tree $\mathbf{T}$, corresponding (on consecutive layers) to mean μ^1_+ and μ^1_-. Then we define $C^+_{(n)} = P(S^+_\emptyset = 1|N_v; |v| \le 2n) = 1 - \prod_{i_1=1}^{N_\emptyset}(1 - \prod_{i_2=1}^{N_{i_1}}(...(1 - (\mu^1_+)^{N_{i_1 i_2 \dots i_{2n-1}}})...))$. It follows that $C^+_{(n)}$ converges a.s. and in L^2 to C^+ and that this must be the unique endogeneous solution (since if Z is any solution with mean μ^1_+ then $E[Z_\emptyset|N_v; |v| \le 2n] = C^+_{(n)}$).

As in Lemma 4.3, we establish that there are at most two solutions by showing that there are at most two possible moment sequences for a solution and that if μ^1_+ is stable (for $f^{(2)}$) then the only possible moment sequence corresponds to the discrete solution S^+.

To do this, note that, denoting a possible moment sequence starting with first moment μ^1_+ by (μ^k_+), we have

$$H(\mu^k_-) = H(\sum_{j=0}^{k}(-1)^j \binom{k}{j} H(\mu^j_+)) = \sum_{j=0}^{k}(-1)^j \binom{k}{j} \mu^j_+. \tag{6.2}$$

Then we look for solutions of

$$H(\sum_{j=0}^{k-1}(-1)^j \binom{k}{j} H(\mu^j_+) + (-1)^k H(t)) = \sum_{j=0}^{k-1}(-1)^j \binom{k}{j} \mu^j_+ + (-1)^k t, \tag{6.3}$$

in the range where the argument of H on the lefthand side is non- negative and less than 1. In this range H is increasing and convex so there are at most two solutions.

Suppose that μ^1_+ is a stable fixed point then the unique moment sequence is constant, since the other solution of

$$g(t) \stackrel{def}{=} H(1 - 2H(\mu^1_+) + H(t)) - (1 - 2\mu^1_+ + t) = 0$$

must be greater than μ^1_+ (because $g'(\mu^1_+) = H'(\mu^1_+)H'(\mu^1_-) - 1 \le 0$).

If μ_+^1 is unstable, then there are, potentially two solutions for μ_+^2, one of which is μ_+^1. Taking the other potential solution, and seeking to solve (6.3), one of the solutions will give a value for μ_-^k greater than $\mu^* > \mu_-^2$ which is not feasible, so there will be at most one sequence with $\mu_+^2 \neq \mu_+^1$.

Now, as in the proof of Theorem 4.9, we can show that, if μ_+^1 is unstable then, in the corresponding RDE with branching factor truncated by n, the two solutions to the RDE are distinct for large n, and the endogenous solution converges to C^+ in L^2 as $n \to \infty$. It follows that there are two distinct solutions in this case.

□

Given a fixed point μ_+^1 of $f^{(2)}$, denote the law of the corresponding conditional probability solution by ς_+. Denote the corresponding basin of attraction (under $\mathcal{T}^2$) by $B(\varsigma_+)$ and denote the basin of attraction of μ_+^1 under the map $f^{(2)}$ by $b^2(\mu_+^1)$. Then

Theorem 6.3. *the following dichotomy holds:*

(i) *if* $H'(\mu_+^1)H'(\mu_-^1) > 1$, *then*

$$B(\varsigma_+) = \{\pi : \pi \text{ has mean } \mu_+^1 \text{ and } \pi \text{ is not concentrated on } \{0,1\}\}.$$

(ii) *if* $H'(\mu_+^1)H'(\mu_-^1) \leq 1$ *then*

$$B(\varsigma_+) = \{\pi : \pi \text{ has mean } m \in b^2(\mu_+^1)\}$$

Proof. This can be proved in exactly the same way as Theorems 5.3 and 5.4. □

7. Examples

We conclude with some examples.

Example 7.1. *We consider first the case where N is Geometric(α), so that $P(N = k) = \beta^{k-1}\alpha$ and $H(s) = \frac{\alpha s}{1-\beta s}$ (with $\beta = 1-\alpha$). It follows that*

$$f^{(2)}(s) = s,$$

so that every pair $(s, \frac{1-s}{1-qs})$ is a two-cycle of f and the unique fixed point of f is $1-\sqrt{\alpha}$. It also follows that s is a neutrally stable fixed point of $f^{(2)}$ for each $s \in [0,1]$. Thus we see that the unique endogenous solution to the original RDE is discrete and the value at the root of the tree is the a.s. limit of $1-\prod_{i_1=1}^{N_\emptyset}(1-\prod_{i_2=1}^{N_{i_1}}(\ldots(1-(1-\sqrt{\alpha})^{N_{i_1,\ldots,i_n}})\ldots))$. Moreover, for any s, there is a unique solution to the iterated RDE with mean s and it is discrete and endogenous and is the a.s. limit of $1-\prod_{i_1=1}^{N_\emptyset}(1-\prod_{i_2=1}^{N_{i_1}}(\ldots(1-s^{N_{i_1,\ldots,i_{2n-1}}})\ldots))$.

Example 7.2. *Consider the original noisy veto-voter model on the binary tree. It follows from (2.3) that*

$$H(z) = (pH(z)+qz)^2 \Rightarrow H(z) = \frac{1-2pqz-\sqrt{1-4pqz}}{2p^2}.$$

This is non-defective if and only if $p \leq \frac{1}{2}$ (naturally), i.e. if and only if extinction is certain in the trimmed tree from the original veto-voter model. It is fairly straightforward to show that $H'(\mu^1) > 1 \Leftrightarrow p < \frac{1}{2}$. Thus, the endogenous solution is non-discrete precisely when the trimmed tree is sub-critical.

Example 7.3. *In contrast to the case of the veto-voter model on the binary tree, the veto-voter model on a trinary tree can show a non-endogenous discrete solution even when the trimmed tree is supercritical. More precisely, the trimmed tree is supercritical precisely when* $p > \frac{1}{3}$*, but the discrete solution is non-endogenous if and only if* $p < p_e^{(3)} \stackrel{def}{=} \frac{3.\sqrt{3}-4}{3.\sqrt{3}-2}$*, and* $p_e^{(3)} > \frac{1}{3}$*.*

References

[1] D. Aldous and A. Bandyopadhyay (2005): A survey of max-type recursive distributional equations, *Ann. Appl. Probab.*, **15**, 2, 1047 - 1110.

[2] A. Bandyopadhyay (2006): A necessary and sufficient condition for the tail-triviality of a recursive tree process, *Sankhya*, **68**, 1, 1–23.

[3] W. Feller (1971): "Introduction to probability theory and its applications", Vol. 2, 2nd Edn., Wiley, New York

[4] L. Rüschendorf (2006): On stochastic recursive equations of sum and max type, *J. Appl. Prob.*, **43**, 687 - 703.

[5] J. Warren (2006): Dynamics and endogeny for processes indexed by trees, *Stochastics*, **78**, 5, 327–342.

Dept. of Statistics, University of Warwick, Coventry CV4 7AL, UK

E-mail address: s.d.jacka@warwick.ac.uk