

Endogeny for Recursive Tree Processes

Antar Bandyopadhyay

(Joint work with Professor David J. Aldous)

IBM T. J. Watson Research Center

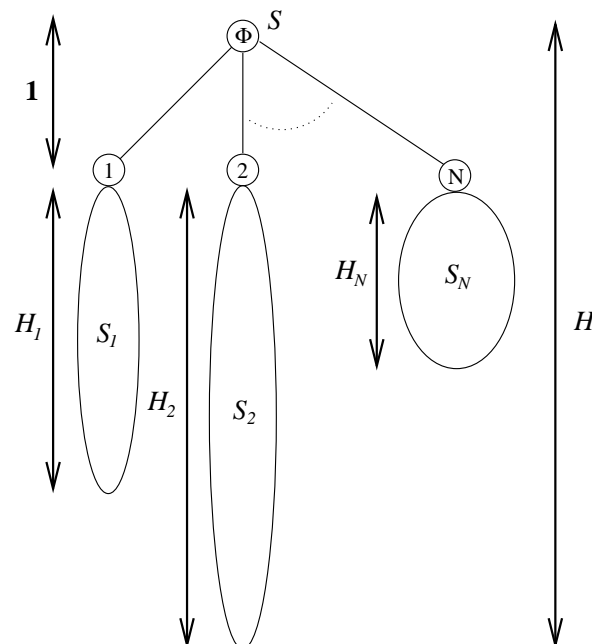
Institute for Mathematics and Its Applications
University of Minnesota

(<http://www.ima.umn.edu/~antar>)

April 30, 2004

Three Examples

Examples 1 : Consider a (*sub*)-critical Galton-Watson branching process with the progeny distribution N , so $\mathbf{E}[N] \leq 1$; we assume $\mathbf{P}(N = 1) < 1$.

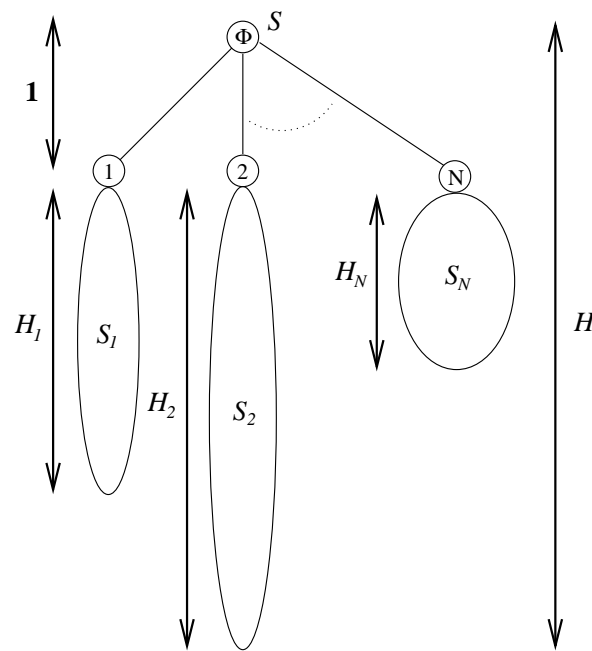


Height of the Tree : Let $H := 1 +$ height of the G-W tree, then $H < \infty$ a.s. and

$$H \stackrel{d}{=} 1 + \max(H_1, H_2, \dots, H_N) \quad \text{on } \mathbb{N},$$

where $(H_j)_{j \geq 1}$ are i.i.d. with same law as of H and are independent of N .

Examples 2 : Consider the same *(sub)-critical* Galton-Watson branching process.



Size of the Tree : Let $S :=$ total size of the tree. Once again $S < \infty$ a.s. since the process is *(sub)-critical*. Further

$$S \stackrel{d}{=} 1 + (S_1 + S_2 + \cdots + S_N) \quad \text{on } \mathbb{N},$$

where $(S_j)_{j \geq 1}$ are i.i.d. with same law as of S and are independent of N .

We will call such equations *Recursive Distributional Equations* (RDE).

Example 3 (Quicksort Algorithm/Distribution) :

- Select the first number from a pile of n numbers and divide the other $(n - 1)$ numbers into two piles, according to *less than* or *bigger* than the first number.
- Recursively sort the two piles (which are now smaller in size).
- $X(n) := \#$ comparisons needed to sort n numbers starting from a uniform random permutation of $[n]$. Then

$$X(n) \stackrel{d}{=} X_1(U_n) + X_2(n - 1 - U_n) + (n - 1),$$

where $X_1(\cdot)$ and $X_2(\cdot)$ are i.i.d. with same law as of $X(\cdot)$ and are independent of U_n which is uniform on $\{0, 1, 2, \dots, n - 1\}$.

- Rösler (1990) showed $\mathbf{E}[X(n)] \sim 2n \log n$ and moreover

$$\frac{X(n) - 2n \log n}{n} \xrightarrow{d} Y,$$

where distribution of Y satisfies the RDE

$$Y \stackrel{d}{=} UY_1 + (1 - U)Y_2 + C(U),$$

where Y_1 and Y_2 are i.i.d. with same law as of Y and are independent of $U \sim \text{Uniform}[0, 1]$, and $c(u) := 1 + 2u \log u + 2(1 - u) \log(1 - u)$.

Typical features of RDEs

$$\text{Ex. 1 : } X \stackrel{d}{=} 1 + \max(X_1, X_2, \dots, X_N) \text{ on } \mathbb{N}$$

$$\text{Ex. 2 : } X \stackrel{d}{=} 1 + (X_1 + X_2 + \dots + X_N) \text{ on } \mathbb{N}$$

$$\text{Ex. 3 : } X \stackrel{d}{=} UX_1 + (1 - U)X_2 + C(U) \text{ on } \mathbb{R}$$

- **Unknown Quantity** : Distribution of X .
- **Known Quantities** :
 - $N \leq \infty$ which may or may not be random (e.g. $N \equiv 2$ in Ex. 3).
 - Possibly some more randomness whose distribution is known (e.g. U in the Ex. 3).
 - How we combine the known and unknown randomness (e.g. “ $1 + \max$ ” operation in Ex. 1).
- **What is the RDE doing ?** To find a distribution μ such that when we take i.i.d. samples $(X_j)_{j \geq 1}$ from it and only use N many of them (where N is independent of the samples) and do the manipulation then we end up with another sample $X \sim \mu$.

Remark : In the case $N = 1$ a.s. it reduces to the question of finding a stationary distribution of a discrete time Markov chain.

Two main uses of RDEs

- **Direct use** : The RDE is used directly to define a distribution. Examples include,
 - ▶ The height (and also the size) of a (sub)-critical Galton-Watson tree (the first two examples).
 - ▶ The Quicksort distribution (Example 3).
 - ▶ Discounted tree sums / inhomogeneous percolation on trees.
 - ▶ ... *and many others*.
- **Indirect use**: The RDE is used to define some auxiliary variables which help in defining/characterizing some other quantity of interest. Among others the following two type of applications are of special interest
 - ▶ 540° *argument* ! (will give an example).
 - ▶ Determining critical points and scaling laws (will not give an example).

General Setup

- Let (S, \mathfrak{G}) be a measurable space, and \mathcal{P} be the collection of all probabilities on (S, \mathfrak{G}) .
- Let (ξ, N) be a pair of random variables such that N takes values in $\{0, 1, 2, \dots; \infty\}$.
- Let $(X_j)_{j \geq 1}$ be **i.i.d** S -valued random variables, which are independent of (ξ, N) .
- $g(\cdot)$ is a S -valued measurable function with appropriate domain.

Recursive Distributional Equation (RDE)

Definition 1 *The following fixed-point equation on \mathcal{P} is called a Recursive Distributional Equation (RDE)*

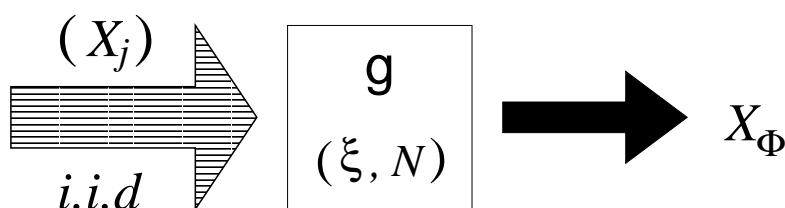
$$X \stackrel{d}{=} g(\xi; X_j, 1 \leq j \leq^* N), \quad \text{on } S$$

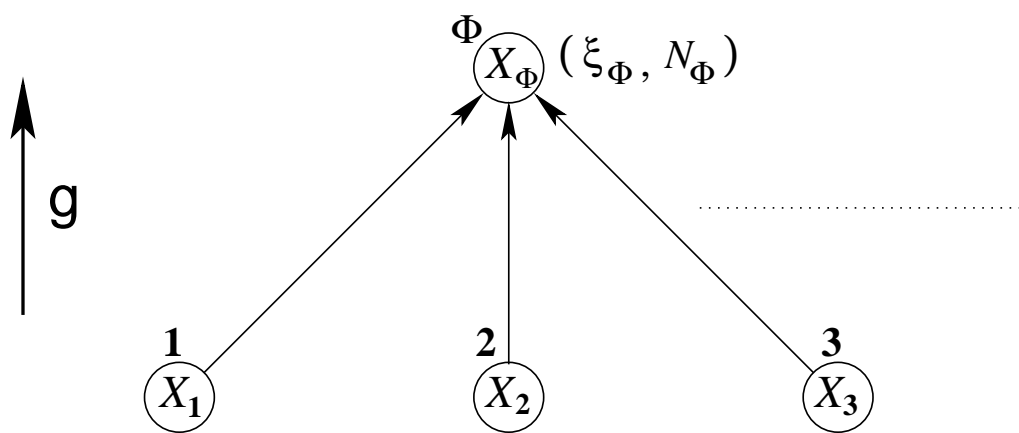
where $(X_j)_{j \geq 1}$ are independent copies of X and are independent of (ξ, N) .

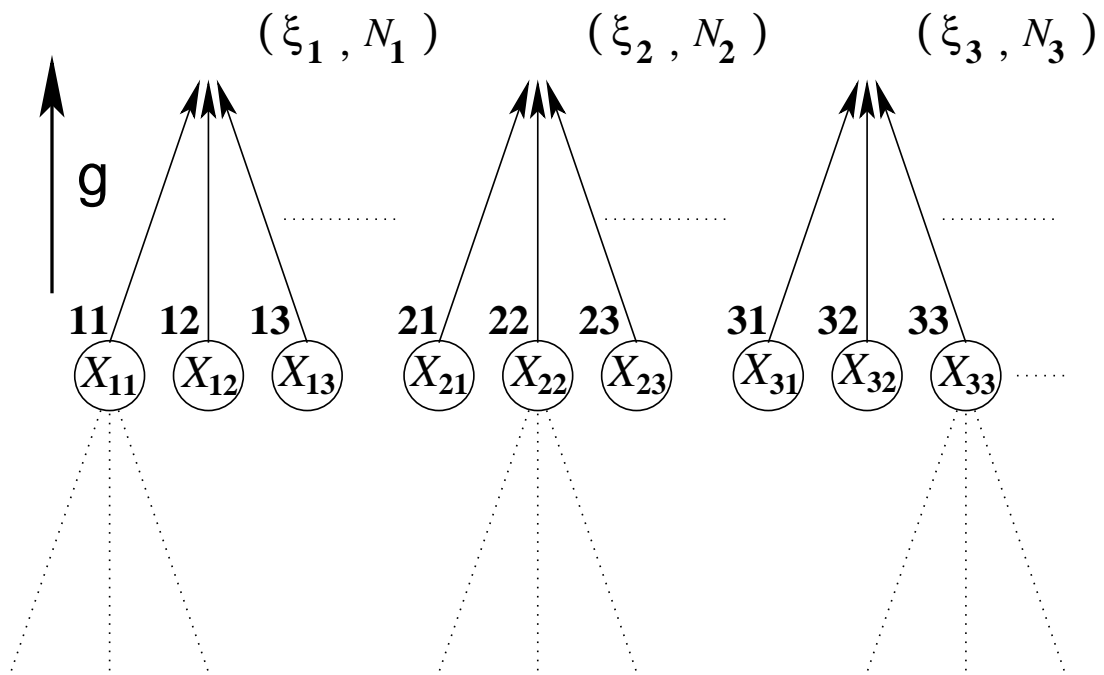
Remark : A more conventional (analysis) way of writing the equation would be

$$\mu = T(\mu)$$

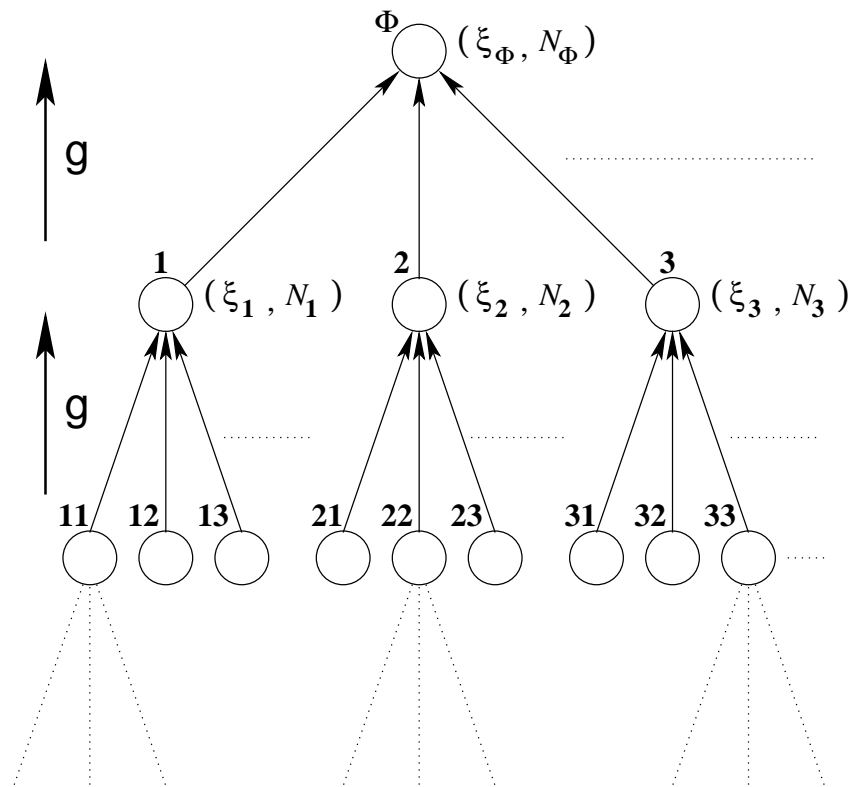
where T is the operator associated with the above equation, which depends on the function g and the joint distribution of the pair (ξ, N) , and μ is the (unknown) law of X .





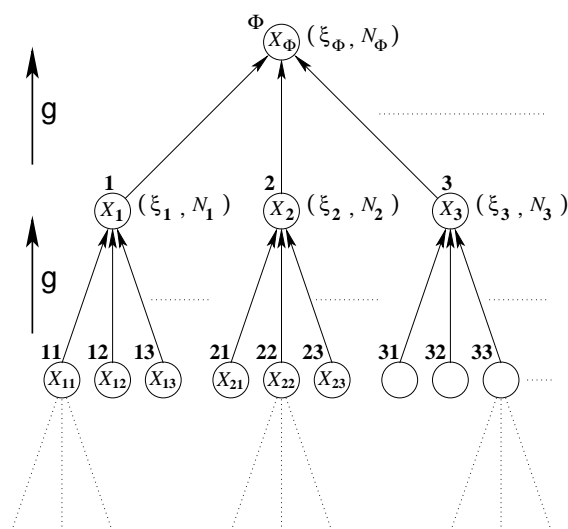


Recursive Tree Framework (RTF)



- **Skeleton** : $\mathbb{T}_\infty := (\mathcal{V}, \mathcal{E})$ is the canonical infinite tree with vertex set $\mathcal{V} := \{\mathbf{i} \mid \mathbf{i} \in \mathbb{N}^d, d \geq 1\} \cup \{\emptyset\}$, and edge set $\mathcal{E} := \{e = (\mathbf{i}, \mathbf{i}j) \mid \mathbf{i} \in \mathcal{V}, j \in \mathbb{N}\}$, and root \emptyset .
- **Innovations** : Collection of i.i.d pairs $\{(\xi_{\mathbf{i}}, N_{\mathbf{i}}) \mid \mathbf{i} \in \mathcal{V}\}$.
- **Function** : The function $g(\cdot)$.

Recursive Tree Process (RTP)



Consider a **RTF** and let μ be a solution of the associated **RDE**. A collection of S -valued random variables $(X_{\mathbf{i}})_{\mathbf{i} \in \mathcal{V}}$ is called an invariant *Recursive Tree Process (RTP)* with marginal μ if

- $X_{\mathbf{i}} \sim \mu \quad \forall \mathbf{i} \in \mathcal{V}$.
- $X_{\mathbf{i}} = g(\xi_{\mathbf{i}}; X_{\mathbf{i}j}, 1 \leq j \leq^* N_{\mathbf{i}}) \quad \forall \mathbf{i} \in \mathcal{V}$.
- $X_{\mathbf{i}}$ is independent of $\{(\xi_{\mathbf{i}'}, N_{\mathbf{i}'}) \mid |\mathbf{i}'| < |\mathbf{i}|\}$, for all $\mathbf{i} \in \mathcal{V} \setminus \{\emptyset\}$, where $|\mathbf{i}| = d$ if $\mathbf{i} \in \mathbb{N}^d$.

Remark : Using *Kolmogorov's consistency*, an invariant RTP with marginal μ exists if and only if μ is a solution of the associated RDE.

Endogeny

Natural Question : Does X_\emptyset only depend on the innovation process (the *data*) $(\xi_i, N_i)_{i \in \mathcal{V}}$?

Definition 2 Let \mathcal{G} be the σ -field generated by the innovation process $\{(\xi_i, N_i) \mid i \in \mathcal{V}\}$. We will say an invariant RTP is endogenous if X_\emptyset is \mathcal{G} -measurable.

Motivations

- Presence / absence of *external* randomness.
- Influence of the boundary at infinity !
- Relation with *long-range independence* ? (recent work of Gamarnik, Nowicki, Swirszcz, 2004)

One easy fact to built our confidence

Remark : Associated with a RTF there is a Galton-Watson branching process tree rooted at \emptyset defined only through $\{N_i | i \in \mathcal{V}\}$, call it \mathcal{T} . Essentially any associated invariant RTP lives on \mathcal{T} .

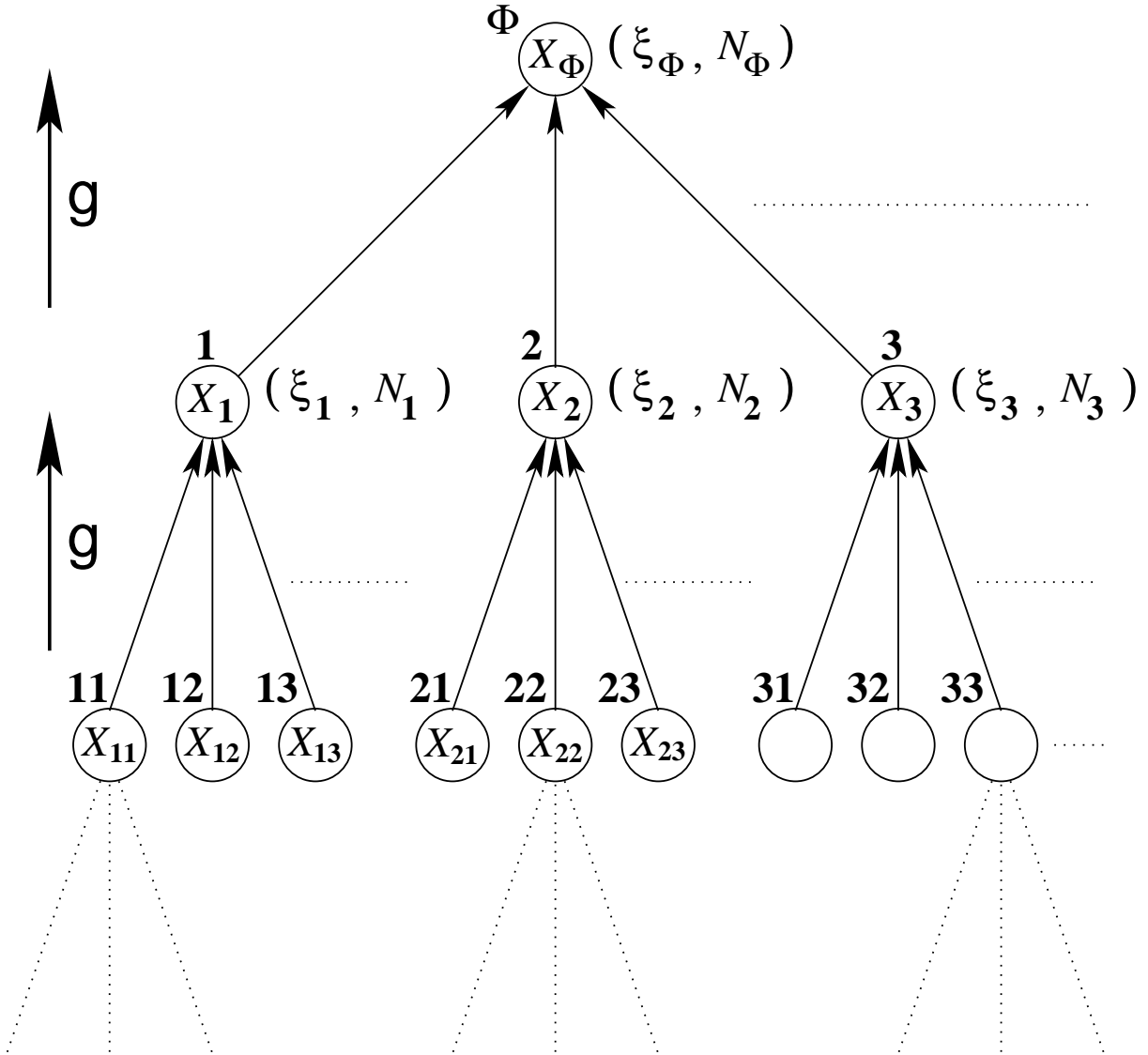
Proposition 1 *If \mathcal{T} is almost surely finite (equivalently $E[N] \leq 1$) then the associated RDE has unique solution and the RTP is endogenous.*

Remarks :

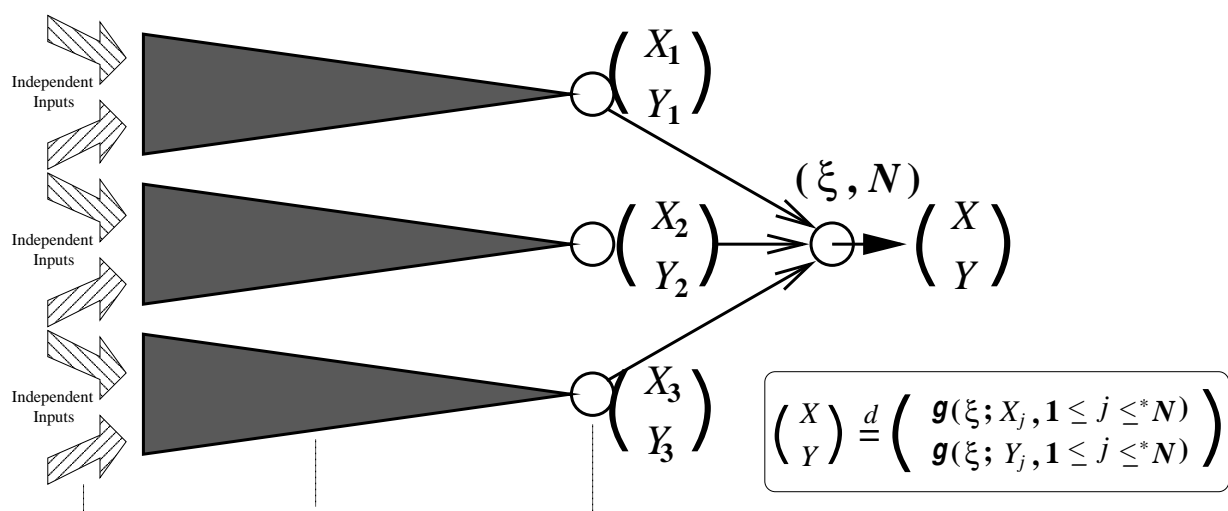
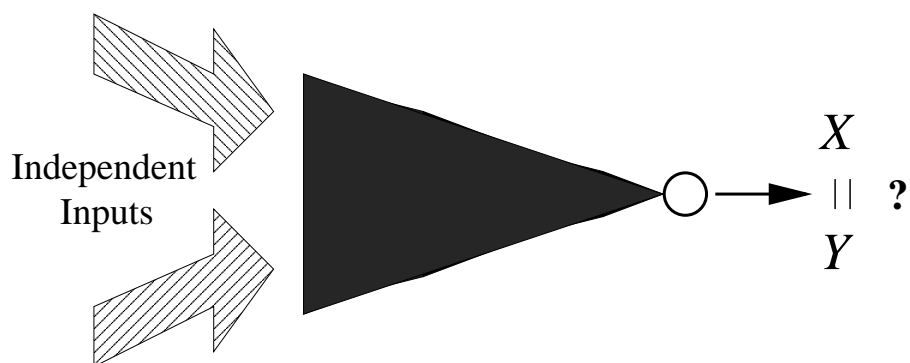
- The RDEs in the first two examples have unique solutions and are endogenous.
- Perhaps the simplest example of a RDE with no non-trivial endogenous solution is the following

$$X \stackrel{d}{=} \frac{X_1 + X_2}{\sqrt{2}}.$$

The solution set is the Normal($0, \sigma^2$) family. But the associated RTF has *no randomness* involved and hence none of the non-trivial RTP is endogenous.



Input at Infinity RTF Output



Bivariate Uniqueness

Consider the following **bivariate RDE**,

$$\begin{pmatrix} X \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} g(\xi; X_j, 1 \leq j \leq^* N) \\ g(\xi; Y_j, 1 \leq j \leq^* N) \end{pmatrix}$$

where $(X_j, Y_j)_{j \geq 1}$ are i.i.d and has the same law as of (X, Y) , and are independent of the innovation (ξ, N) .

Definition 3 *An invariant RTP with marginal μ has **bivariate uniqueness** property if the above bivariate RDE has unique solution as $X = Y$ a.s on the space of joint probabilities with both marginals μ .*

An Equivalence Theorem

Theorem 1 *Suppose the S is a Polish space. Consider an invariant RTP with marginal distribution μ .*

(a) If the endogenous property holds then the bivariate uniqueness property holds.

(b) Conversely, (under some technical conditions) if the bivariate uniqueness property holds and then the endogenous property holds.

(c) If $T^{(2)}$ be the operator associated with the bivariate RDE then endogenous property holds if and only if

$$T^{(2)n}(\mu \otimes \mu) \xrightarrow{d} \mu^{\nearrow},$$

where $\mu \otimes \mu$ is the product measure, and μ^{\nearrow} is the measure concentrated on the diagonal with both marginal μ .

Remark : Results of similar type can also be found in the study of Gibbs measures and Markov random fields.

Successful Use and/or Application of Endogeny

- **Characterization** : Some time one can show that only the “*fundamental*” solution of an RDE is endogenous. For example one can show that for the *Quicksort RDE* only the limiting *Quicksort* distribution is endogenous.
- **540° argument** :
 - ▶ Can construct *approximate* solution for the random assignment problem by using endogenous optimal solution of the matching problem on **PWIT**. (will not discuss this application in this talk)
 - ▶ Can show existence of an automorphism invariant version of *frozen percolation* process on an infinite regular tree without having presence of any external randomness.

Frozen Percolation on Regular Binary Tree

The Setup :

- Let $\mathbb{T}_3 = (\mathbb{V}, \mathbb{E})$ be the infinite regular binary tree.
- Each edge $e \in \mathbb{E}$ is equipped with independent edge weight $U_e \sim \text{Uniform}[0, 1]$.
- Think of time moving from 0 to 1.

Frozen Percolation Process (informal description):

- For an edge $e \in \mathbb{E}$ at the time instance $t = U_e$ open the edge e if each of its end vertex is in a finite component; otherwise do not open e .
- Let $(\mathcal{A}_t)_{t \geq 0}$ be set process of open edges starting from $\mathcal{A}_0 = \emptyset$.

The Regular Percolation Process :

- For an edge $e \in \mathbb{E}$ at the time instance $t = U_e$ open the edge e .
- If $(\mathcal{B}_t)_{t \geq 0}$ be the set process of open edges the it can be described as

$$\mathcal{B}_t = \{e \in \mathbb{E} \mid U_e \leq t\}$$

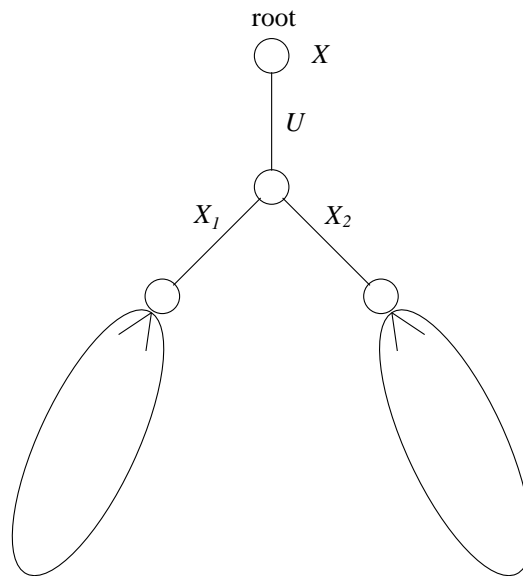
Remarks : Unlike the regular percolation process it is not clear whether the *frozen percolation process* exists and if so whether it admits a simpler description using only the edge weights.

Two Easy Observations : If frozen percolation process exists then following must hold

- $\mathcal{A}_t \subseteq \mathcal{B}_t$ for all $t \in [0, 1]$.
- $\mathcal{A}_t = \mathcal{B}_t$ if $t \leq \frac{1}{2}$ (since the critical probability for infinite binary tree is $\frac{1}{2}$).

540° Argument [Aldous, 2000]

- **Stage 1** : Suppose that the process exists on \mathbb{T}_3 . Let $\widetilde{\mathbb{T}}_3$ be the *planted* binary tree which is a modification of \mathbb{T}_3 where we distinguish a vertex of degree 1 as the *root* and all other vertices have degree 3.



- ▶ $X :=$ Time it takes for the root to join ∞ (will write $X = \infty$ if it never joins).
- ▶ $X_j :=$ Time it takes for the root to join to ∞ in the j^{th} sub-tree for $j = 1, 2$.
- ▶ X_1 and X_2 are independent copies of X .
- ▶ It is easy to see that

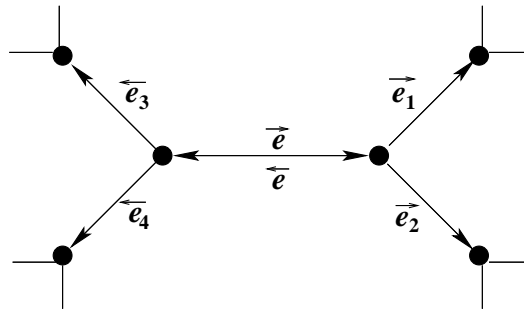
$$X \stackrel{d}{=} \begin{cases} X_1 \wedge X_2 & \text{if } X_1 \wedge X_2 > U \\ \infty & \text{otherwise} \end{cases}$$

- **Stage 2 :**

- ▶ The RDE has only one solution with full support given by

$$\mu(dy) = \frac{dy}{2y^2}, \quad \frac{1}{2} < y < 1, \quad \mu(\{\infty\}) = \frac{1}{2}.$$

So using the general theory we can construct the invariant RTP with marginal μ .



- ▶ Each edge $e \in \mathbb{E}$ defines two directed edges, and each directed edge \vec{e} defines one *planted tree*, let $X_{\vec{e}}$ be the corresponding X variable.
- ▶ Each directed edge \vec{e} has two children say \vec{e}_1 and \vec{e}_2 then $\{X_{\vec{e}_1}, X_{\vec{e}_2}\}$ and $X_{\vec{e}}$ satisfies the equation with the edge weight U_e .
- ▶ Each edge $e \in \mathbb{E}$ has a set of four *children* which are the four directed edges away from e . We denote it by $\partial\{e\}$.
- ▶ Define $\mathcal{A}_1 := \{e \in \mathbb{E} \mid U_e < \min(X_f : f \in \partial\{e\})\}$ and $\mathcal{A}_t := \{e \in \mathcal{A}_1 \mid U_e \leq t\}$ for $0 \leq t < 1$.

- **Stage 3** : Using this *external* random variables $(X_{\vec{e}})$ repeat the original computation to prove the existence of a frozen percolation process on \mathbb{T}_3 . In fact it is easy to see that this construction gives an automorphism invariant version of the process.

Remark :

- The construction of the process not only uses the edge weights (U_e) but also (possibly) *external* random variables, namely $(X_{\vec{e}})$.
- If we can prove that the solution μ of the frozen percolation RDE is endogenous then it will automatically follow that the variables $(X_{\vec{e}})$ are measurable with respect to the edge weights (U_e) . Thus the process (\mathcal{A}_t) as constructed above will not have any *external randomness*. This will then imply that the informal description defines a process on \mathbb{T}_3

Frozen Percolation RDE

- Recall the RDE associated with the frozen percolation process,

$$X \stackrel{d}{=} \Phi(X_1 \wedge X_2; U)$$

where X_1, X_2 are independent copies of X and are independent of $U \sim \text{Uniform}[0, 1]$ and the function Φ is given by

$$\Phi(x; u) := \begin{cases} x & \text{if } x > u \\ \infty & \text{otherwise} \end{cases} .$$

- Also recall that it has *unique* solution with full support given by

$$\mu(dy) = \frac{dy}{2y^2}, \quad \frac{1}{2} < y < 1, \quad \mu(\{\infty\}) = \frac{1}{2}.$$

Theorem 2 *The invariant RTP with marginal μ has bivariate uniqueness property, that is, the following bivariate RDE has unique solution as $X = Y$ a.s with marginal μ*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \Phi(X_1 \wedge X_2; U) \\ \Phi(Y_1 \wedge Y_2; U) \end{pmatrix}$$

where $(X_j, Y_j)_{j=1,2}$ are independent copies of (X, Y) , and are independent of $U \sim \text{Uniform}[0, 1]$.

Corollary 2.1 *The invariant RTP with marginal μ is endogenous. Thus the frozen percolation process on \mathbb{T}_3 as constructed is measurable with respect to the edge weights.*

Outline of the proof of Theorem 2

- Notice that X and Y have the same distribution μ . So if $F(x, y) = \mathbf{P}(X \leq x, Y \leq y)$ and $G(x, y) = \mathbf{P}(X > x, Y > y)$ then for every $x, y \in [\frac{1}{2}, 1]$

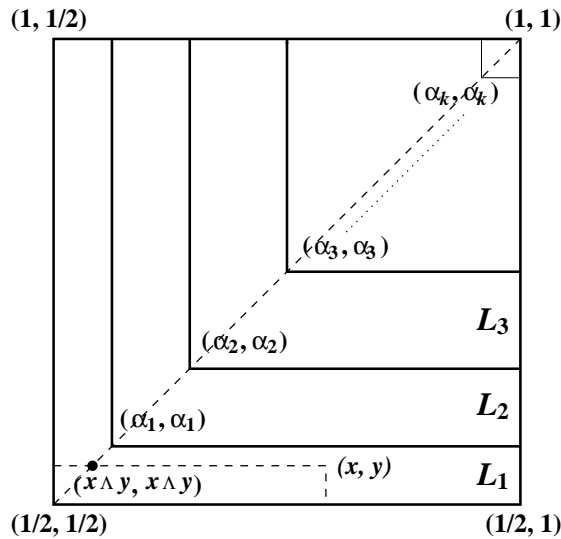
$$G(x, y) = F(x, y) + \frac{1}{2x} + \frac{1}{2y} - 1.$$

- From the bivariate RDE we get

$$F(x, y) =$$

$$\int_0^{x \wedge y} (G^2(x, y) - G^2(x, u) - G^2(u, y) + G^2(u, u)) du.$$

- We know that $X = Y$ a.s. is a solution so $G_0(x, y) = \frac{1}{2(x \vee y)}$ is a solution of the integral equation. It is enough to prove that $G = G_0$ is the *only* solution.
- Let $H(x, y) = 1 - \frac{G(x, y)}{G_0(x, y)}$, so we need to show $H \equiv 0$ on $D := [\frac{1}{2}, 1]^2$.



- Substituting back into the equation and after some algebra we get

$$H(x, y) = \frac{1}{G_0(x, y)} \int_0^{x \wedge y} \Lambda(x, y, u) du,$$

where Λ is a function (has long expression !) which satisfies the estimate

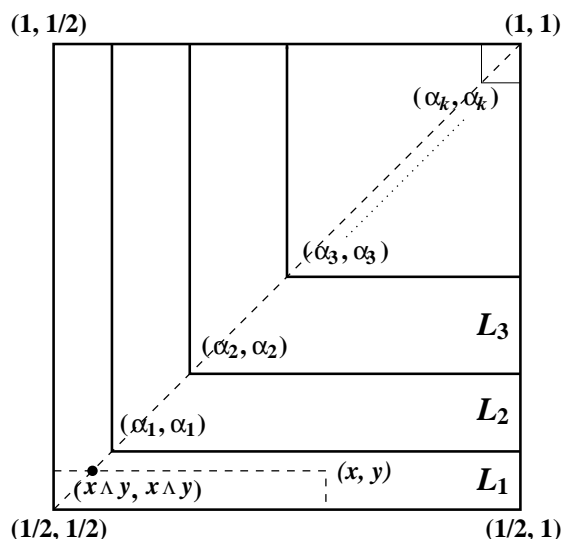
$$|\Lambda(x, y, u)| \leq 4G_0^2(u, u) (2|H(x, y)| + |H^2(x, y)|),$$

whenever $u \leq x \wedge y$.

- Find $\frac{1}{2} = \alpha_0 < \alpha_1 < \dots < \alpha_k = 1$ such that

$$\int_{\alpha_{i-1}}^{\alpha_i} G_0^2(u, u) du < \frac{1}{48}.$$

- We partition D into L -shape parts (as in the figure) where $L_i := \{(x, y) \mid \alpha_{i-1} \leq x \wedge y < \alpha_i\}$.



- Define $\| H \|_i := \sup_{x,y \in L_i} |H(x, y)|$.
- Start with $i = 1$, let $(x, y) \in L_1$. Note $G_0(x, y) \geq \frac{1}{2}$. Thus from the estimate of Λ we get

$$\begin{aligned}
 |H(x, y)| &\leq 24 \| H \|_i \int_{\alpha_{i-1}}^{\alpha_i} G_0^2(u, u) du \\
 &\leq \frac{1}{2} \| H \|_i
 \end{aligned}$$

- Thus $H \equiv 0$ on L_1 . Now proceed inductively for $i = 2, 3, \dots, k$ to conclude $H \equiv 0$ on D .