Electron. Commun. Probab. **22** (2017), no. 20, 1–16.
DOI: 10.1214/17-ECP47
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Galton-Watson probability contraction

Moumanti Podder[*] Joel Spencer[†]

Abstract

We are concerned with exploring the probabilities of first order statements for Galton-Watson trees with $Poisson(c)$ offspring distribution. Fixing a positive integer k, we exploit the k-move Ehrenfeucht game on rooted trees for this purpose. Let Σ, indexed by $1 \leq j \leq m$, denote the finite set of equivalence classes arising out of this game, and D the set of all probability distributions over Σ. Let $x_j(c)$ denote the true probability of the class $j \in \Sigma$ under $Poisson(c)$ regime, and $\vec{x}(c)$ the true probability vector over all the equivalence classes. Then we are able to define a natural recursion function Γ, and a map $\Psi = \Psi_c : D \to D$ such that $\vec{x}(c)$ is a fixed point of Ψ_c, and starting with any distribution $\vec{x} \in D$, we converge to this fixed point via Ψ because it is a contraction. We show this both for $c \leq 1$ and $c > 1$, though the techniques for these two ranges are quite different.

Keywords: Galton-Watson trees; almost sure theory; first order logic.
AMS MSC 2010: Primary 05C20, Secondary 60F20.
Submitted to ECP on February 19, 2016, final version accepted on January 23, 2017.

1 Introduction

The Poisson Galton-Watson tree (henceforth, GW tree) $T = T_c$ with parameter $c > 0$ is a much studied object. It is a random rooted tree. Each node, independently, has Z children where Z has the Poisson distribution with mean c. We let P_c denote the probability under T_c.

We shall examine the first order language on rooted trees. This consists of a constant symbol R (the root), equality $v = w$, and a parent function $\pi(v)$ defined for all vertices $v \neq R$. (Purists may prefer a binary relation $\pi^*[v, w]$, that w is the parent of v.) Sentences must be finite and made up of the usual Boolean connectives $(\neg, \vee, \wedge, \ldots)$ and existential $\exists_v$ and universal $\forall_v$ quantification over vertices. The *quantifier depth* of a sentence A is the depth of the nesting of the existential and universal quantifiers. A formal definition of quantifier depth may be found in Section 1.2 of [5], along with many other aspects of first order logic on random structures.

Example 1.1. No node has precisely one child.

$$\neg[\exists_{u,x} \pi(x) = u \wedge \forall_z [(z \neq x) \Rightarrow \neg\pi(z) = u]]. \tag{1.1}$$

[*]Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, United States. E-mail: mp3460@nyu.edu

[†]Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, United States. E-mail: spencer@cims.nyu.edu

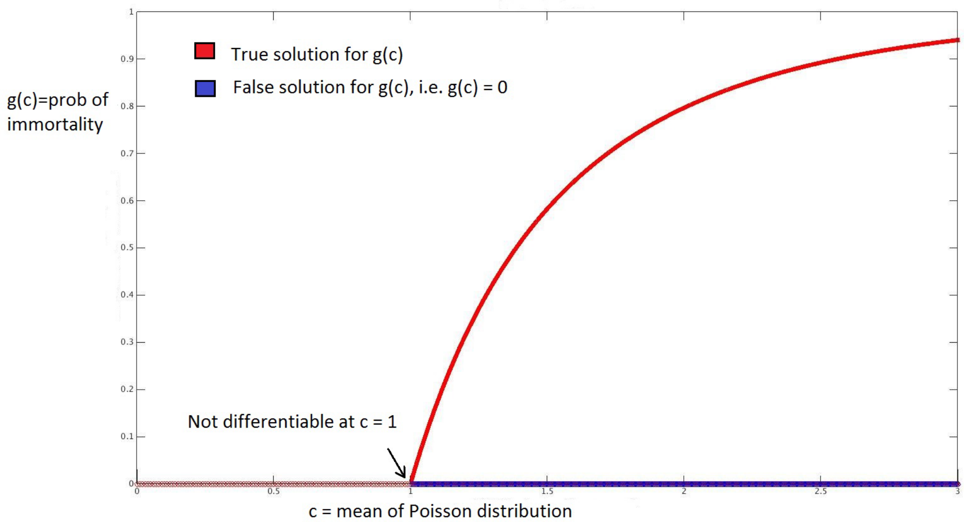
We outline some of our results. For any first order sentence A set

$$f_A(c) = \Pr[T_c \models A], \tag{1.2}$$

the probability that $T = T_c$ has property A. Except in examples we will work with the *quantifier depth* k of A. The value k shall be arbitrary but *fixed* throughout this presentation. We will show in (2.3) that there is a finite dimensional probability vector $\vec{x} = \vec{x}(c)$ that depends only on c (in a smooth manner), and that each f_A is the sum of a subset of the coordinates of x. In Section 2.4 we show that the $x_j(c)$ are solutions to a finite system of equations involving polynomials and exponentials. The solution is described as the fixed point of a map Ψ_c over the m-dimensional simplex D, where m is the number of different sentences (up to logical equivalence which is made more precise later) of quantifier depth k. In Theorem 6.1 we show that this system has a unique solution. In Sections 3.3 (for the subcritical case) and 7 (for the general case) we show that Ψ_c is a contraction. Employing the Implicit Function Theorem in Section 8 we then achieve one of our main results:

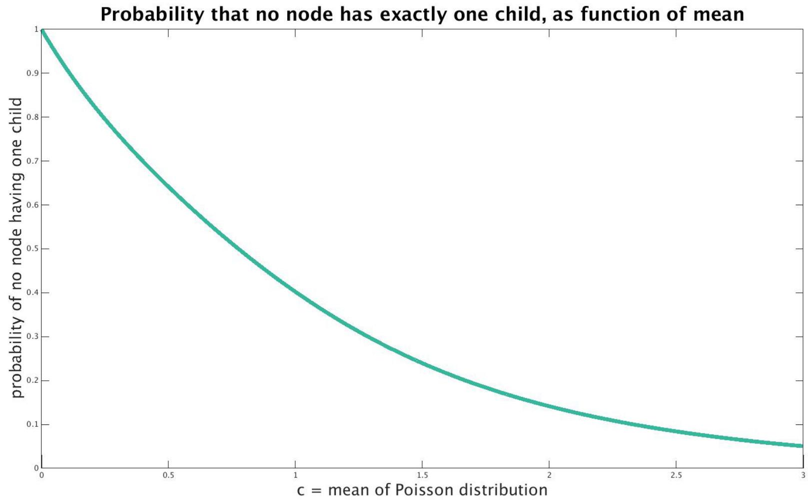
Theorem 1.2. *Let A be first order. Then $f_A(c)$ is a $C^\infty(0,\infty)$ function. That is, all derivatives of $f_A(c)$ exist and are continuous at all $c > 0$.*

Remark 1.3. Let $y = g(c)$ be the probability that $T = T_c$ is infinite. It is well known that $g(c) = 0$ for $c \leq 1$ while for $c > 1$, $y = g(c)$ is the unique positive real satisfying $e^{-cy} = 1 - y$. The value $c = 1$ is often referred to as a critical, or percolation, point for GW-trees. The function $g(c)$ is *not* differentiable at $c = 1$. The right sided derivative $\lim_{c\to 1^+}(g(c) - g(1))/(c-1)$ is 2 while the left sided derivative is zero. An interpretation of Theorem 1.2 that we favor is that the critical point $c = 1$ cannot be seen through a First Order lens. Theorem 1.2 thus yields that the property of T being infinite is not expressible in the first order language – though this can be shown with a much weaker hammer!



The plot clearly shows how the function is not differentiable at $c = 1$, and how the solution is non-unique for $c > 1$.

But the plot corresponding to the property in Example 1.1 shows that the probability is a smooth function of c, which is in keeping with Theorem 1.2.



Definition 1.4. *With $v \in T$, $T(v)$ denotes the rooted tree consisting of v and all of its descendants, with v regarded as the root. For s a non-negative integer, $T|_s$ denotes the rooted tree consisting of its vertices up to generation at most s. We call $T|_s$ the s-cutoff of T. (This is defined even if no vertices are at generation s.) $T(v)|_s$ denotes the s-cutoff of $T(v)$.*

2 The Ehrenfeucht game

2.1 Ehrenfeucht game

The Ehrenfeucht game is a well-known tool used to analyze first order properties on rooted trees. It serves as a bridge between mathematical logic and a complete structural description of logical statements on rooted trees (and other structures). In order to describe the game, we fix an arbitrary positive integer k, and two trees T_1 rooted at R_1 and T_2 rooted at R_2. The game, denoted by $EHR[T_1, T_2, k]$, is played between two players, the Spoiler and the Duplicator, consisting of k rounds. In each round Spoiler picks a vertex from *either* T_1 or T_2 and then Duplicator picks a vertex from the other tree. Letting (x_i, y_i) denote the pair chosen from T_1, T_2 respectively in the i-th round for $1 \leq i \leq k$, Duplicator wins if *all* of the following hold:

1. $x_i = R_1$ iff $y_i = R_2$;
2. $\pi(x_j) = x_i$ iff $\pi(y_j) = y_i$, i.e. x_i is the parent of x_j if and only if y_i is the parent of y_j;
3. $\pi(x_i) = R_1$ iff $\pi(y_i) = R_2$, i.e., if x_i is a child of the root R_1, then y_i is a child of R_2, and vice versa;
4. $x_i = x_j$ iff $y_i = y_j$.

2.2 Equivalence classes

Let k denote an arbitrary positive integer. We may then define an equivalence relation $\equiv_k$ (we often omit the subscript) on all trees T, as follows.

Definition 2.1. *$T \equiv_k T'$ if they have the same truth value for all first order A of quantifier depth at most k.*

Remark 2.2. The critical significance of the Ehrenfeucht games is that $T \equiv_k T'$ if and only if Duplicator wins the k-move Ehrenfeucht game $EHR[T, T', k]$.

Let $\Sigma = \Sigma_k$ denotes the set of all equivalence classes defined by this relation. Critically, Σ_k is a *finite* set. As a function of k we note that $|\Sigma_k|$ grows like a tower function. We give Chapter 2 of [5], Section 6.1 of [1], and Lemma 2.4.8 of [3] as a general reference to these basic results. In particular, Exercise 6.11 of [1] shows that indeed the number of equivalence classes resulting out of $\equiv_k$ is finite for every fixed but arbitrary k.

Definition 2.3. *For any rooted tree T, the Ehrenfeucht value of T, denoted by $EV[T]$, is that equivalence class $\sigma \in \Sigma$ to which T belongs.*

For convenience we denote the elements of Σ by $\Sigma = \{1, \ldots, m\}$. We let $D \subset \mathbb{R}^m$ denote the set of all possible probability distributions over Σ. That is,

$$D = \left\{ (x_1, \ldots, x_m) : \sum_{j=1}^{m} x_j = 1 \text{ and all } x_j \geq 0 \right\}. \tag{2.1}$$

Recall that P_c denotes the probability under T_c, the GW process with $Poisson(c)$ offspring distribution. Let

$$x_j(c) = P_c(j) = P[EV[T_c] = j], \quad j \in \Sigma. \tag{2.2}$$

Then $\vec{x}(c) = (x_j(c) : 1 \leq j \leq m)$ denotes the probability vector in D under P_c. That is, $\vec{x}(c)$ denotes the probability distribution for the equivalence classes in case of $T = T_c$.

Theorem 2.4. *$\vec{x}(c)$ has derivatives of all orders. In particular, each $x_j(c)$ has derivatives of all orders.*

The proof of Theorem 2.4 is a goal of this paper, accomplished only in Section 8 after many preliminaries. Any first order sentence A of quantifier depth $\leq k$ is determined, tautologically, by the set $S(A)$ of those $j \in \Sigma$ such that all T with $EV[T] = j$ have property A. For any $j \in \Sigma$ either all T with $EV[T] = j$ or no T with $EV[T] = j$ have property A. We may therefore decompose the $f_A(c)$ of (1.2) into

$$f_A(c) = \sum_{j \in S(A)} x_j(c). \tag{2.3}$$

Theorem 1.2 will therefore follow from Theorem 2.4.

2.3 Recursive states

In the k-move Ehrenfeucht game, values $\geq k$ are roughly all "the same." This will be made precise in the subsequent Lemma 2.5. We define

$$C = \{0, 1, \ldots, k-1, \omega\}. \tag{2.4}$$

The phrase "there are ω copies" is to be interpreted as "there are $\geq k$ copies." We call $v \in T$ a *rootchild* if its parent is the root R. For $w \neq R$ we say v is the *rootancestor* of w if v is that unique rootchild with $w \in T(v)$. Of course, a rootchild is its own rootancestor.

Lemma 2.5 roughly states that the Ehrenfeucht value of a tree T is determined by the Ehrenfeucht values $EV[T(v)]$ for all the rootchildren v. To clarify: ω rootchildren means at least k rootchildren while n rootchildren, $n \in C$, $n \neq \omega$ means precisely n rootchildren.

Lemma 2.5. *Let $\vec{n} = (n_1, \ldots, n_m)$ with all $n_j \in C$. Let T have the property that for all $1 \leq j \leq m$ there are n_j rootchildren v with $EV[T(v)] = j$. Then $\sigma = EV[T]$ is uniquely determined.*

Definition 2.6. *Let*

$$\Gamma : \{(n_1, \ldots, n_m) : n_i \in C \text{ for all } 1 \leq i \leq m\} \to \Sigma \tag{2.5}$$

be given by $\sigma = \Gamma(\vec{n})$ with $\vec{n}, \sigma$ satisfying the conditions of Lemma 2.5. Then Γ is called the recursion function.

Proof of Lemma 2.5. Let T, T' have the same $\vec{n}$. We give a strategy for Duplicator in the Ehrenfeucht game $EHR[T, T', k]$. Duplicator will create a partial matching between the rootchildren $v \in T$ and the rootchildren $v' \in T'$. When v, v' are matched, $EV[T(v)] = EV[T'(v')]$. At the end of any round of the game call a rootchild $v \in T$ (similarly $v' \in T'$) *free* if no $w \in T(v)$ (correspondingly no $w' \in T'(v')$) has yet been selected.

Suppose Spoiler plays $w \in T$ (similarly $w' \in T'$) with rootancestor v. Suppose v is free. Duplicator finds a free $v' \in T'$ with $EV[T(v)] = EV[T'(v')]$. When $EV[T(v)] = j \in \Sigma$ and $n_j \neq \omega$, then as the number of rootchildren of T with Ehrenfeucht value j is exactly the same as that in T', hence this can be done. In the special case where $n_j = \omega$ the vertex v' may be found as there have been at most $k-1$ moves prior to this move and so there are at most $k-1$ rootchildren $w' \in T'$ with $EV[T'(w')] = j$ that are not free. Duplicator then matches v, v'. Duplicator can win $EHR[T(v), T'(v'), k]$ as $EV[T(v)] = EV[T'(v')]$. Once v, v' have been matched any move $z \in T(v)$ is responded to with a move $z' \in T'(v')$, and vice versa, using the strategy for $EHR[T(v), T'(v'), k]$. □

Remark 2.7. Tree automata consist of a finite state space Σ, an integer $k \geq 1$, a map Γ as in (2.5) and a notion of accepted states. While first order sentences yield tree automata, the notion of tree automata is broader. Tree automata roughly correspond to second order monadic sentences, a topic we hope to explore in future work.

2.4 Solution as fixed point

We come now to the central idea. We define, for $c > 0$, a map $\Psi_c : D \to D$. Let $\vec{x} = (x_1, \ldots, x_m) \in D$, a probability distribution over Σ. Imagine root R has Poisson mean c children. To each child we assign, independently, a $j \in \Sigma$ with distribution $\vec{x}$. Let $n_j \in C$ be the number of children assigned j. Let $\vec{n} = (n_1, \ldots, n_m)$. Apply the recursion function (equation 2.5) Γ to get $\sigma = \Gamma(\vec{n})$. We then define $\Psi_c(\vec{x})$ to be the distribution of this random σ.

The special nature of the Poisson distribution allows a concise expression. When the initial distribution is $\vec{x}$, the number of children assigned j will have a Poisson distribution with mean cx_j *and* these numbers are mutually independent over $j \in \Sigma$. Thus

$$\Pr[n_j = u] = e^{-cx_j} \frac{(cx_j)^u}{u!} \quad \text{for } u \in C, u \neq \omega, \tag{2.6}$$

and

$$\Pr[n_j = \omega] = 1 - \sum_{u=0}^{k-1} \Pr[n_j = u]. \tag{2.7}$$

From the independence, for any $\vec{a} = (a_1, \ldots, a_m)$ with $a_1, \ldots, a_m \in C$,

$$\Pr[\vec{n} = \vec{a}] = \prod_{j=1}^{m} \Pr[n_j = a_j]. \tag{2.8}$$

Thus, writing $\Psi_c(x_1, \ldots, x_m) = (y_1, \ldots, y_m)$,

$$y_j = \Sigma \Pr[\vec{n} = \vec{a}] \tag{2.9}$$

where the summation is over all $\vec{a}$ with $\Gamma(\vec{a}) = j$.

We place all Ψ_c into a single map Δ:

$$\Delta : D \times (0, \infty) \to D \text{ by } \Delta(\vec{x}, c) = \Psi_c(\vec{x}). \tag{2.10}$$

Setting $\Delta(x_1, \ldots, x_m, c) = (y_1, \ldots, y_m)$, the y_j are finite sums of products of polynomials and exponentials in the variables $x_1, \ldots, x_m, c$. In particular, all partial derivatives of all orders exist everywhere.

Recall (2.2), $\vec{x}(c)$ denotes the probability distribution for the equivalence classes under the probability measure P_c for $T = T_c$.

Lemma 2.8. *Let $\Psi_c : D \to D$ be as defined in 2.4, where D is the set of all probability distributions on Σ, the set of equivalence classes from the k-round Ehrenfeucht game in the first order setting, as defined in 2.1. Suppose $\vec{x}(c)$ is a fixed point for $\Psi_c : D \to D$. That is, $\Psi_c(\vec{x}(c)) = \vec{x}(c)$.*

Proof. By definition of Γ, we know, for any $j \in \Sigma$,

$$\begin{aligned}\Psi_c(\vec{x}(c))(j) &= \sum_{\vec{a}\in C^{\Sigma}:\Gamma(\vec{a})=j} \prod_{i=1}^{m} P[n_i = a_i | \vec{x} = \vec{x}(c)], \quad \text{from (2.8)},\\ &= \sum_{\vec{a}\in C^{\Sigma}:\Gamma(\vec{a})=j} \prod_{i=1}^{m} P[Poisson(c \cdot x_i(c)) = a_i]\\ &= \sum_{\vec{a}\in C^{\Sigma}:\Gamma(\vec{a})=j} P_c[\vec{n} = \vec{a}]\\ &= P_c[j] = x_j(c),\end{aligned}$$

where recall that P_c is the probability induced under the $Poisson(c)$ offspring distribution. □

Example 2.9. For many particular A the size of Σ, which may be thought of as the state space, may be reduced considerably. Let A be the property given in (1.1), that no node has precisely one child. For every node in a given tree, we define state 1 which denotes that A is true for the subtree rooted at this node, and state 2 which denotes that A is false for the subtree rooted at this node. We set $C = \{0, 1, \omega\}$ with ω meaning "at least two." Let $n_1, n_2 \in C$ be the number of rootchildren v with $T(v)$ having state $1, 2$ respectively. Then T is in state 1 if and only if $\vec{n} = (n_1, n_2)$ has one of the values $(0, 0), (\omega, 0)$. Let D be the set of distributions on the two states, $D = \{(x, y) : 0 \le x \le 1, y = 1 - x\}$. Then $\Psi_c(x, y) = (z, w)$ with $w = 1 - z$ and

$$z = e^{-cx}e^{-cy} + e^{-cy}[1 - e^{-cx} - (cx)e^{-cx}] = e^{-cy}[1 - (cx)e^{-cx}]. \tag{2.11}$$

The fixed point (x, y) then has $x = \Pr[A]$ satisfying the equation

$$x = e^{-c(1-x)}[1 - (cx)e^{-cx}]. \tag{2.12}$$

Example 2.10. Let A be that there is a vertex v with precisely one child who has precisely one child. Let state 1 be that A is true. Let state 2 be that A is false but that the root has precisely one child. Let state 3 be all else. Set $C = \{0, 1, \omega\}$, with ω meaning ≥ 2. Set $D = \{x, y, z : x + y + z = 1, x \ge 0, y \ge 0, z \ge 0\}$. T is in state 1 if and only if $\vec{n} = (n_1, n_2, n_3)$ has either $n_1 \neq 0$ or $n_1 = 0, n_2 = 1, n_3 = 0$. T is in state 2 if and only if $\vec{n} = (0, 0, 1)$. Then $\vec{x}(c) = (x, y, z)$ must satisfy the system (noting $z = 1 - x - y$ is dependent)

$$x = (1 - e^{-cx}) + e^{-cx}(cye^{-cy})e^{-cz} = 1 - e^{-cx} + cye^{-c} \tag{2.13}$$

$$y = e^{-cx}e^{-cy}(cze^{-cz}) = cze^{-c} \tag{2.14}$$

Here $x = \Pr[A]$.

In general, however, $\Pr[A]$ will be the sum (2.3).

3 The contraction formulation

3.1 The total variation metric

On D we let $||x-y||_2$ denote the usual Euclidean metric, and $||\vec{x}-\vec{y}||_1$ the L^1 distance. We let $TV(\vec{x},\vec{y})$ denote the total variation distance. With $\vec{x}=(x_1,\ldots,x_m)$ and $\vec{y}=(y_1,\ldots,y_m)$ this standard metric is given by

$$TV(\vec{x},\vec{y})=\frac{1}{2}||\vec{x}-\vec{y}||_1=\frac{1}{2}\sum_{j=1}^{m}|x_j-y_j|. \tag{3.1}$$

Total variation distance between any two probability distributions μ,ν on the same probability space has a natural interpretation in terms of coupling μ and ν. This is captured in the following standard result:

Theorem 3.1. *For any two probability distributions μ,ν on a common probability space,*

$$TV(\mu,\nu)=\min\{P[X\neq Y]:(X,Y)\text{ any coupling of }\mu,\nu\}. \tag{3.2}$$

The coupling at which this minimum is indeed attained is known as the *optimal* coupling. We refer the reader to Chapter 4 of [2] for further reading, especially Proposition 4.7 for a proof of existence of the optimal coupling.

3.2 The contraction theorem

Theorem 3.2. *Fix any arbitrary $c>0$. Let $\Psi_c:D\to D$ be as defined in 2.4, where D is the set of all probability distributions on Σ, the set of equivalence classes from the k-round Ehrenfeucht game in the first order setting, as defined in 2.1. Then there exists a positive integer s and an $\alpha<1$ such that for all $\vec{x},\vec{y}\in D$,*

$$||\Psi_c^s(\vec{x})-\Psi_c^s(\vec{y})||_2\leq\alpha||\vec{x}-\vec{y}||_2. \tag{3.3}$$

The proof of this result is given in several parts over the next few sections. The subcritical case is considered in Subsection 3.3, whereas the proof in the supercritical regime requires a more delicate analysis, which is given in detail in Section 7.

The map $\Psi_c^s:D\to D$ has a natural interpretation. Let $\vec{x}=(x_1,\ldots,x_m)\in D$. Generate a random GW tree $T=T_c$ but stop at generation s (the root is at generation 0). To each node (there may not be any) v at generation s assign independently $j\in\Sigma$ from distribution $\vec{x}$. Now we work up the tree towards the root. Suppose, formally by induction, that the nodes w at generation i have been assigned some $j\in\Sigma$. A v at generation $i-1$ will then have n_j children assigned j (allowing $n_j=\omega$). The value at v, which is now determined by Lemma 2.5, is given by the recursion function $\Gamma(\vec{n})$ of Definition 2.5. $\Psi_c^s(\vec{x})$ will then be the distribution of the Ehrenfeucht value assigned to the root.

Remark 3.3. While this work examines first order properties on rooted trees, it is instructive to cosider the non-first order property A that T is infinite. Set $C=\{0,\omega\}$ (ω denoting ≥ 1) and let state 1 be that T is infinite, state 2 that it is not. T is in state 1 if and only if $\vec{n}=(\omega,0)$ or (ω,ω). Then $D=\{(x,1-x):0\leq x\leq 1\}$ and

$$\Psi_c(x,1-x)=(1-e^{-cx},e^{-cx}). \tag{3.4}$$

However, when $c>1$, Ψ_c has two fixed points: $(0,1)$ and the "correct" $(x(c),1-x(c))$. The contraction property (3.3) will not hold. With ϵ small, $1-e^{-c\epsilon}\sim c\epsilon$ and so $\vec{x}=(0,1)$, and $\vec{y}=(\epsilon,1-\epsilon)$ become further apart on application of Ψ_c, when $c>1$.

3.3 The subcritical case

Here we prove Theorem 3.2 under the additional assumption that $c < 1$. The proof in this case is considerably simpler. Further, it may shed light on the general proof.

Theorem 3.4. *Fix any arbitrary $c < 1$. Let $\Psi_c : D \to D$ be as defined in 2.4, where D is the set of all probability distributions on Σ, the set of equivalence classes from the k-round Ehrenfeucht game in the first order setting, as defined in 2.1. Then, for any $\vec{x}, \vec{y} \in D$,*

$$TV(\Psi_c(\vec{x}), \Psi_c(\vec{y})) \leq c \cdot TV(\vec{x}, \vec{y}). \tag{3.5}$$

Proof. The main idea is to use suitable coupling of $\vec{x}$ and $\vec{y}$. First we fix $s \in \mathbb{N}$. We then create two pictures. In both pictures, let v have s many children $v_1, \ldots v_s$. In picture 1, we assign, mutually independently, labels $X_1, \ldots X_s \in \Sigma$ to $v_1, \ldots v_s$ respectively, with $X_i \sim \vec{x},\ 1 \leq i \leq s$. In picture 2, we assign, again mutually independently, labels $Y_1, \ldots Y_s \in \Sigma$ to $v_1, \ldots v_s$, with $Y_i \sim \vec{y},\ 1 \leq i \leq s$. The pairs $(X_i, Y_i),\ 1 \leq i \leq s$, are mutually independent, but for every i, (X_i, Y_i) is coupled so that

$$P[X_i \neq Y_i] = TV(\vec{x}, \vec{y}). \tag{3.6}$$

Suppose X_v is the label of the root v in picture 1 that we get from the recursion function Γ (from (2.5)), and Y_v that in picture 2. Then $X_v \sim \Psi_c(\vec{x}), \quad Y_v \sim \Psi_c(\vec{y})$.

$$\begin{aligned} TV(\Psi_c(\vec{x}), \Psi_c(\vec{y})) &\leq P[X_v \neq Y_v] \\ &\leq \sum_{s=0}^{\infty} P[Poisson(c) = s] \sum_{i=1}^{s} P[X_i \neq Y_i] \\ &= \sum_{s=0}^{\infty} P[Poisson(c) = s] \cdot s \cdot TV(\vec{x}, \vec{y}) \\ &= c \cdot TV(\vec{x}, \vec{y}). \end{aligned}$$

□

Lemma 3.5. *Fix any $c < 1$. Let $\Psi_c : D \to D$ be as defined in 2.4, where D is the set of all probability distributions on Σ, the set of equivalence classes from the k-round Ehrenfeucht game in the first order setting, as defined in 2.1. Then the conclusion of Theorem 3.2 holds.*

Proof. The inequalities

$$||\vec{z}||_1 \geq ||\vec{z}||_2 \geq m^{-1/2} ||\vec{z}||_1 \tag{3.7}$$

bound the L^1 and L^2 norms on $\mathbb{R}^m$ by multiples of each other. As $TV(\vec{x}, \vec{y}) = \frac{1}{2}||\vec{x} - \vec{y}||_1$,

$$TV(\vec{x}, \vec{y}) \geq \frac{1}{2}||\vec{x} - \vec{y}||_2 \geq m^{-1/2} \cdot TV(\vec{x}, \vec{y}). \tag{3.8}$$

Applying Theorem 3.4 repeatedly,

$$TV(\Psi_c^s(\vec{x}), \Psi_c^s(\vec{y})) \leq c^s \cdot TV(\vec{x}, \vec{y}). \tag{3.9}$$

Combining (3.8 and 3.9)

$$||\Psi_c^s(\vec{x}) - \Psi_c^s(\vec{y})||_2 \leq 2 \cdot TV(\Psi_c^s(\vec{x}), \Psi_c^s(\vec{y})) \leq 2c^s \cdot TV(\vec{x}, \vec{y}) \leq 2c^s \sqrt{m}\ ||\vec{x} - \vec{y}||_2. \tag{3.10}$$

We select s so that $2c^s \sqrt{m} < 1$ and set $\alpha = 2c^s \sqrt{m}$. □

4 Universality

We define a function $Rad[i]$ on the nonnegative integers by the recursion

$$Rad[0] = 0 \text{ and } Rad[i+1] = 3Rad[i] + 1 \text{ for } i \geq 0. \tag{4.1}$$

Definition 4.1. *In T we define a distance $\rho(v, w)$ to be the minimal r for which there is a sequence $v = z_0, z_1, \ldots, z_r = w$ where each z_{i+1} is either the parent* or *a child of z_i. In other words, this is really the ordinary graph distance. We set $\rho(v, v) = 0$.*

As an example, cousins would be at distance four.

Definition 4.2. *For r a nonnegative integer, $v \in T$, the ball of radius r around v, denoted $B(v, r)$ is the set of $w \in T$ with $\rho(v, w) \leq r$. We consider v a designated vertex of $B(v, r)$.*

We define an equivalence relation, depending on k, on such balls.

Definition 4.3. *We define $B(v, r) \equiv_k B(v', r)$ if the two sets satisfy the same first order sentences of quantifier depth at most $k-1$ with v, v' as designated vertices, allowing the relations π and $=$.*

This is the same notation as used in Definition 2.1.

Remark 4.4. Note that the $(k-1)$-round Ehrenfeucht game with v, v' designated is identical to the k-round Ehrenfeucht game in which the first round is mandated to select v, v'.

Equivalently, $B(v, r) \equiv_k B(v', r)$ if Duplicator wins the k-move Ehrenfeucht game on these sets in which the first round is mandated to be v, v' and Duplicator must preserve the relations π and $=$. Let Σ_k^{BALL} denote the set of all equivalence classes under the equivalence relation $\equiv_k$, on all such balls as defined in 4.2.

Definition 4.5. *We say $S_1, S_2 \subset T$ are* strongly disjoint *if there are no $v_1 \in S_1$, $v_2 \in S_2$ with $\rho(v_1, v_2) \leq 1$.*

Definition 4.6. *We say T is k*-full *if for any $v_1, \ldots, v_{k-1} \in T$ and any $\sigma \in \Sigma_k^{BALL}$ there exists a vertex v such that*

1. *$B(v, Rad[k-1])$ is in equivalence class σ.*
2. *$B(v, Rad[k-1])$ is strongly disjoint from all $B(v_i, Rad[k-1])$.*
3. *$B(v, Rad[k-1])$ is strongly disjoint from $B(R, Rad[k])$, R the root.*

Definition 4.7. *Fix a positive integer k. Consider two trees T and T', and two subsets of vertices C in T and C' in T'. Let the induced subgraphs on C and C' both be connected. Suppose we are also given nodes $z_1, \ldots z_w$ in C and $z'_1, \ldots z'_w$ in C' for some $1 \leq w \leq k$. We say that C and C' are k-equivalent if the Duplicator can win the k-move Ehrenfeucht game on it, maintaining $=$ and π, where z_i and z'_i, $1 \leq i \leq w$, are designated vertices.*

Remark 4.8. Note that the k-move Ehrenfeucht game with designated vertices $z_1, \ldots, z_w \in T$, $z_1 \ldots, z'_w \in T$ is equivalent to the $k+w$-move Ehrenfeucht game in which the first w moves are mandated to be $z_i \in T, z'_i \in T'$

When T is k-full our next result shows that the truth value of first order sentences of quantifier depth at most k is determined by examining T "near" the root. This "inside-outside" strategy is well known, see, for example, [5].

Theorem 4.9. *Let T, T' with roots R, R' both be k-full. Suppose, as per Definition 4.3, $B(R, Rad[k]) \equiv_{k+1} B(R', Rad[k])$. Then T, T' have the same k-Ehrenfeucht value as given by Definition 2.3.*

Proof. Let T, T' satisfy the conditions of Theorem 4.9. We give a strategy for Duplicator to win the k-move Ehrenfeucht game. For convenience we add a round zero in which the roots R, R' are selected. Duplicator will insure inductively that with i moves remaining the following somewhat complicated condition, denoted as $OK[i]$, holds:

Let $x_0, x_1, \ldots, x_{k-i} \in T$, $y_0, y_1 \ldots, y_{k-i} \in T'$, where x_0 and y_0 are the roots of T and T' respectively. Let S_i denote the union of the balls $B(x_j, Rad[i])$, $0 \le j \le k-i$ and let S'_i denote the union of the balls $B(y_j, Rad[i])$, $0 \le j \le k-i$. The set S_i splits T and the set S'_i splits T' into components. Duplicator wants to ensure that for every such component C in T and the corresponding component C' in T', if $x_j \in C$ then $y_j \in C'$ and vice versa; also, C and C' should be equivalent components in the sense of Definition 4.7, with exactly those x_j (correspondingly y_j), $0 \le j \le k-i$, as designated vertices that are in C (correspondingly C').

When $i = k$, $OK[k]$ holds because $x_0 = R$, $y_0 = R'$ and by the hypothesis of Theorem 4.9, $B(R, Rad[k]) \equiv_{k+1} B(R', Rad[k])$.

Now suppose, by induction, that $k \ge i > 0$ and $x_0, x_1, \ldots, x_{k-i} \in T$, $y_0, y_1 \ldots, y_{k-i} \in T'$ satisfy $OK[i]$, with x_0 the root of T and y_0 that of T'. In the next round suppose (the other case being identical) Spoiler selects $x_{k-i+1} \in T$. Now there are two possible cases:

1. *Inside case:* There exists x_j with $0 \le j \le k-i$, such that $\rho(x_j, x_{k-i+1}) \le 2Rad[i-1]+1$. Then the ball $B(x_{k-i+1}, Rad[i-1])$ is entirely contained in $B(x_j, Rad[i])$, from the recursion (4.1)). So, we can choose the corresponding v in $B(y_j, Rad[i])$ which allows him to win the game on the components containing $B(x_j, Rad[i])$, $B(y_j, Rad[i])$, of k rounds. By induction hypothesis, the components formed by $\bigcup_{l=0}^{k-i} B(x_l, Rad[i])$ are equivalent; the balls $B(x_{k-i+1}, Rad[i-1])$ and $B(y_{k-i+1}, Rad[i-1])$ are contained in the balls $B(x_j, Rad[i])$ and $B(y_j, Rad[i])$ respectively. Consequently, the components formed by $\bigcup_{l=0}^{k-i+1} B(x_l, Rad[i-1])$ will also be equivalent.
2. *Outside case:* Suppose x_{k-i+1} is chosen such that $\rho(x_j, x_{k-i+1}) > 2Rad[i-1]+1$ for all $0 \le j \le k-i$. Using the definition of k-full as in 4.6, we know that there exists $v \in T'$ such that $B(v, Rad[k-1])$ is strongly disjoint from $B(y_j, Rad[k-1])$ for all $1 \le j \le k-i$, and from $B(R', Rad[k-1])$; further, $B(v, Rad[k-1]) \equiv_k B(x_{k-i+1}, Rad[k-1])$. Duplicator chooses $y_{k-i+1} = v$. Then observe that for every $0 \le j \le k-i$, $B(x_j, Rad[i-1])$ (correspondingly $B(y_j, Rad[i-1])$) is strongly disjoint from $B(x_{k-i+1}, Rad[i-1])$ (correspondingly $B(y_{k-i+1}, Rad[i-1])$, and thus in the component $\left\{\bigcup_{l=0}^{k-i} B(x_l, Rad[i-1])\right\}^c$ (correspondingly $\left\{\bigcup_{l=0}^{k-i} B(y_l, Rad[i-1])\right\}^c$); further $B(y_{k-i+1}, Rad[k-1]) \equiv_k B(x_{k-i+1}, Rad[k-1])$.

So, regardless of Spoiler's choice of $x_{k-i+1} \in T$, Duplicator has selected $y_{k-i+1} \in T'$ and $OK[i-1]$ is satisfied.

Thus Duplicator can play the entire k rounds so that at the end of the game $OK[0]$ holds. But the union of the balls of radius zero around the vertices selected are equivalent in T, T'. Hence Duplicator has won! □

Definition 4.10. *We replace the complex notion of k-full by a simpler sufficient condition. For each $\sigma \in \Sigma_k^{BALL}$ create k copies of a ball in that class. Take a root vertex v and on it place $k \cdot |\Sigma_k^{BALL}|$ disjoint paths (parent to child) of length $Rad[k] + Rad[k-1] + 1$. Identify each endpoint with the roof of one of these copies. We let $UNIV_k$ denote this tree, which we can picture, rather fancifully, as a* Christmas tree*.*

Definition 4.11. *T is called s-universal (given a fixed positive integer k) if all T' with $T|_s \cong T'|_s$ have the same k-Ehrenfeucht value. Thus $EV[T]$ is determined by $T|_s$ completely.*

Theorem 4.12. *If for some v, $T(v) \cong UNIV_k$ then T is k-full. Thus, by Theorem 4.9, the k-Ehrenfeucht value of T is determined by $B(R, Rad[k])$ or $T|_{Rad[k]}$. In other words, T is $Rad[k]$-universal.*

Remark 4.13. Many other trees could be used in place of $UNIV_k$, we use this particular one only for specificity.

Remark 4.14. A subtree $T(v)$, where v is not the root, cannot determine the Ehrenfeucht value of T as, for example, it cannot tell us if the root has, say, precisely two children. Containing this universal tree $UNIV_k$ tells us everything about the Ehrenfeucht value of T *except* properties relating to the local neighborhood of the root.

5 Rapidly determined properties

We consider the underlying probability space for the GW tree $T = T_c$ to be an infinite sequence $X_1, X_2, \dots$ of independent variables, each Poisson with mean c. These naturally create a tree. Let the root have X_1 children. Now we go through the nodes in a breadth first manner. Let the i-th node (the root is the first node) have X_i children. This creates a unique rooted tree. Note, however, that when the tree is finite with, say, n nodes, then the values X_j with $j > n$ are irrelevant. In that case we say that the process *aborts* at time n.

We employ a useful notation of Donald Knuth.

Definition 5.1. *We say an event occurs* quite surely *if the probability that it does not occur drops exponentially in the given parameter.*

Definition 5.2. *Let A be any property or function of rooted trees. We say that A is* rapidly determined *if quite surely (in s, with $T = T_c$ and c given) $X_1, \dots X_s$ tautologically determine A.*

Remark 5.3. Consider the property that T is infinite and suppose $c > 1$. Given $X_1, \dots, X_s$ if the tree has stopped then we know it is finite. Suppose however (as holds with positive limiting probability) after $X_1, \dots X_s$ the tree is continuing. If at that stage there are many nodes we can be reasonably certain that T will be infinite, but we cannot be tautologically sure. This property is *not* rapidly determined.

Remark 5.4. In this work we restrict the language in which A is expressed. It has been suggested that another approach would be to restrict A to rapidly determined properties.

Lemma 5.5. *[4] Let T_0 be an arbitrary finite tree. Let A be the (non first order) property that* either *the process has aborted by time s or there exists $v \in T$ with $T(v) \cong T_0$. Then A is rapidly determined in parameter s.*

The proof is given in [4]. Let T_0 have depth d. Roughly speaking, when we examine $X_1, \dots, X_s$ either the process has aborted or it has not. If not, quite surely some $i \le s\epsilon$ has $T(i) \cong T_0$. Here ϵ is chosen small enough (dependent on c, d) so that quite surely the descendants of all $i \le s\epsilon$, down to d generations, have indices $\le s$.

Lemma 5.6. *Every first order property A is rapidly determined.*

Proof. Let A have quantifier depth k. Let T_0 be the universal tree $UNIV_k$ as given by Definition 4.10. From Lemma 5.5 if T has not aborted by time s then quite surely some $T(i) \cong T_0$. But then T is already k-full and already has depth at least $Rad[k]$. By Theorem 4.9 the k-Ehrenfeucht value of T, hence the truth value of A, is determined solely by $T|_{Rad[k]}$, and hence tautologically by $X_1, \dots X_s$. □

Lemma 5.7. *Fix a positive integer k. Let $T \sim T_c$. Then quite surely (in s), T is s-universal.*

Proof. Lemma 5.6 gives that the k-Ehrenfeucht value of T is quite surely determined by $X_1, \ldots, X_s$. When this is so it is tautologically determined by $T|_s$, which has more information. □

6 Unique fixed point

Theorem 6.1. *Let $\Psi_c : D \to D$ be as defined in 2.4, where D is the set of all probability distributions on Σ, the set of equivalence classes from the k-round Ehrenfeucht game in the first order setting, as defined in 2.1. Then Ψ_c has a* unique *fixed point.*

Proof. Let $f(s)$ be the probability that T_c is *not* s-universal. For any $\vec{y}, \vec{z} \in D$ we couple $\Psi_c^s(\vec{y}), \Psi_c^s(\vec{z})$. Create T_c down to generation s and then give each node at generation s a $\sigma \in \Sigma$, mutually independently, with marginal distribution $\vec{y}$, respectively $\vec{z}$. Then $\Psi_c^s(\vec{y}),\ \Psi_c^s(\vec{z})$ will be the distributions of the induced states of the root. But when T_c is s-universal this will be the same for any $\vec{y}, \vec{z}$. Hence $TV[\Psi_c^s(\vec{y}), \Psi_c^s(\vec{z})] \le f(s)$. When $\vec{y}, \vec{z}$ are fixed points of Ψ, $TV[\vec{y}, \vec{z}] \le f(s)$. As $f(s) \to 0$, $\vec{y} = \vec{z}$. □

Remark 6.2. Theorem 6.1 will also follow from the more powerful Theorem 3.2.

Remark 6.3. It is a challenging exercise to show directly that the solution x to (2.12) or the solution x, y to the system (2.13 and 2.14) are unique.

7 A proof of contraction

7.1 A two stage process

Here we prove Theorem 3.2 for arbitrary c. Let D_0 be the depth of $UNIV_k$, as given by Definition 4.10. We shall set

$$s = s_0 + D_0 \text{ with } s_0 \ge 2 \cdot Rad[k] \tag{7.1}$$

and think of $T|_s$ as being generated in two stages. In Stage 1 we generate $T|_{s_0}$. From Lemma 5.7, by taking s_0 large, this will be s_0-universal with probability near one. In Stage 2 we begin with an arbitrary but fixed T_0 of depth at most s_0. (We say "at most" because it includes the possibility that T_0 has no vertices at depth s_0.) From each node at depth s_0, mutually independently, we generate a GW-tree down to depth D_0. We denote by $Ext(T_0)$ this random tree, now of depth (at most) s.

Definition 7.1. *For any T_0 of depth at most s_0, $BAD[T_0]$ is the event that $Ext(T_0)$ is* not *s_0-universal.*

Lemma 7.2. *There exists positive β such that for any T_0 of depth at most s_0*

$$\Pr(BAD[T_0]) \le e^{-t\beta} \tag{7.2}$$

where t denotes the number of nodes of T_0 at depth s_0.

Proof. Let $v_1, \ldots, v_t$ denote the nodes of T_0 at generation s_0. Each of them independently generates a GW tree. Let $1 - e^{-\beta}$ denote the probability that $T(v_i) \cong UNIV_k$. With probability $e^{-t\beta}$ no $T(v_i) \cong UNIV_k$. But otherwise $Ext(T_0)$ is s_0-universal. □

7.2 Splitting the extension

Let T_0 be an arbitrary tree of depth s_0. Let $\vec{x} \in D$. Assign to the depth s nodes of $Ext(T_0)$ independent identically distributed labels $j \in \Sigma$ taken from distribution $\vec{x}$. Applying the recursion function Γ of Definition 2.6 repeatedly up the generations yields a unique Ehrenfeucht value for the root R.

Definition 7.3. *Let $\Psi_c^s(T_0, \vec{x})$ denote the induced distribution of the Ehrenfeucht value for the root R as derived in the description above.*

Lemma 7.4. *Let $\Psi_c : D \to D$ be as defined in 2.4, where D is the set of all probability distributions on Σ, the set of equivalence classes from the k-round Ehrenfeucht game in the first order setting, as defined in 2.1. Then,*

$$TV(\Psi_c^s(\vec{x}), \Psi_c^s(\vec{y})) \leq \sum \Pr(T|_{s_0} \cong T_0) \cdot TV(\Psi_c^s(T_0, \vec{x}), \Psi_c^s(T_0, \vec{y})) \tag{7.3}$$

where the sum is over all T_0 of depth (at most) s_0.

Proof. We split the distribution of T into the distribution of $Ext(T_0)$, with probability $\Pr[T|_{s_0} \cong T_0]$, over each T_0 of depth (at most) s_0. □

7.3 Some technical lemmas

Let $X = X(c, s)$ be the number of descendants at generation s of the GW tree $T = T_c$. Let Y be the sum of t independent copies of X. The next result (not the best possible) is that the tail of Y is bounded by exponential decay in t.

Lemma 7.5. *There exists $\beta > 0$ and y_0 such that for $y \geq y_0$*

$$\Pr[Y \geq yt] \leq e^{-yt\beta}. \tag{7.4}$$

Proof. Set $f(\lambda) = \ln[E[e^{X\lambda}]]$. We employ Chernoff bounds sub-optimally, taking simply $\lambda = 1$. (We require here a standard argument that $E[e^X]$ is finite.) Then $E[e^Y] = e^{t \cdot f(1)}$ and

$$\Pr[Y \geq yt] \leq E[e^Y]e^{-yt} \leq e^{(f(1)-y)t}. \tag{7.5}$$

For $y \geq 2f(1)$, $f(1) - y \leq -y/2$ and we may take $\beta = \frac{1}{2}$. □

Lemma 7.6. *Let $K, \gamma > 0$. Let BAD be an event with $Pr[BAD] \leq Ke^{-t\gamma}$. Then, for positive constants k, κ,*

$$E[Y \ \mathbf{1}_{BAD}] \leq kte^{-t\kappa}, \tag{7.6}$$

where Y is as defined above.

Remark 7.7. The idea is that, in the worst case, the event BAD would coincide with the event $\{Y \geq s\}$ where $P\,[Y \geq s] = Ke^{-t\gamma}$, for some $s > 0$. But as seen in the following proof, in this situation too, $E[Y \ \mathbf{1}_{BAD}]$ can be bounded as in (7.6).

Proof of Lemma 7.6. We split Y into $\{Y < y_1 t\}$ and $\{Y \geq y_1 t\}$, where y_1 needs to be chosen suitably.

$$E[Y \ \mathbf{1}_{BAD}] = E[Y \ \mathbf{1}_{BAD} \ \mathbf{1}_{Y < y_1 t}] + E[Y \ \mathbf{1}_{BAD} \ \mathbf{1}_{Y \geq y_1 t}], \tag{7.7}$$

where the first term is bounded by $y_1 t P[BAD] \leq K y_1 t e^{-t\gamma}$. The second term is bounded above by $E[Y \ \mathbf{1}_{Y \geq y_1 t}]$, which we bound using Chernoff type arguments. First, recall that Y is the sum of t i.i.d. copies of $X = X(c, s)$ which is the number of nodes at generation s of T_c. Suppose φ_s denotes the cumulant-generating function of X, defined as

$$\varphi_s(\lambda) = \log E[e^{\lambda X}]. \tag{7.8}$$

We fix some $\lambda > 1$ and choose

$$y_1 = \max\left\{y_0, \frac{\gamma + \varphi_s(\lambda)}{\lambda}\right\}, \tag{7.9}$$

where γ is as in the bound of $P[BAD]$, and y_0 as in Lemma 7.5. From (7.4) and (7.9), we then have

$$\begin{aligned} E[Y\ \mathbf{1}_{Y\geq y_1 t}] \leq & y_1 t P[Y\geq y_1 t] + \sum_{k=\lfloor y_1 t\rfloor+1}^{\infty} P[Y\geq k] \\ \leq & y_1 t e^{-y_1 t\beta} + E\left[e^{\lambda Y}\right]\int_{\lfloor y_1 t\rfloor}^{\infty} e^{-\lambda x}dx \\ = & y_1 t e^{-y_1 t\beta} + e^{\varphi_s(\lambda)t}\frac{1}{\lambda}e^{-\lambda\lfloor y_1 t\rfloor} \\ \leq & y_1 t e^{-y_1 t\beta} + e^{\varphi_s(\lambda)t}\frac{e^{\lambda}}{\lambda}e^{-\lambda y_1 t} \\ \leq & y_1 t e^{-y_1 t\beta} + \frac{e^{\lambda}}{\lambda}e^{-\gamma t}. \end{aligned}$$

The desired bound now follows easily, by choosing $\kappa = \min\{y_1\beta,\gamma\}$ and $k = Ky_1 + y_1 + e^{\lambda}/\lambda$.

□

7.4 Bounding expansion

Lemma 7.8. *There exists K_0 (dependent only on s_0, k) such that for* any *T_0 and any $\vec{x},\vec{y}\in D$*

$$TV(\Psi_c^s(T_0,\vec{x}),\Psi_c^s(T_0,\vec{y}))\leq K_0\cdot TV(\vec{x},\vec{y}), \tag{7.10}$$

where $\Psi_c^s(T_0,\vec{x})$ is as defined in Definition 7.3.

Remark 7.9. As K_0 may be large, Lemma 7.8, by itself, does not give a contracting mapping. It does limit how expanding $\Psi_c^s(T_0,\cdot)$ can be.

Remark 7.10. Let t be the number of nodes of T_0 at depth s_0. The expected number of nodes in $Ext(T_0)$ at level $s = s_0 + D_0$ is then tK_1 with $K_1 = c^{D_0}$. The methods of Theorem 3.4 would then give Lemma 7.8 with $K_0 = K_1 t$. However, when $c > 1$ this K_0 would be unbounded in t. Our concern is then with large t, though technically, the proof below works for all t.

Proof. Let t be the number of nodes of T_0 at depth s_0. Let $TV(\vec{x},\vec{y}) = \epsilon$. We again couple $\vec{x},\vec{y}$. As before, let Y be the number of nodes in $Ext(T_0)$ at level s. Given $Y = y$, let us name these nodes $u_1,\dots u_y$. Again we create two pictures. In picture 1, we assign, mutually independently, labels $X_i\in\Sigma$ to u_i, with $X_i\sim\vec{x}$, and in picture 2, label $Z_i\sim\vec{y}$. $(X_i,Z_i), 1\leq i\leq y$ mutually independent, but X_i, Z_i are coupled so that

$$P[X_i\neq Z_i] = TV(\vec{x},\vec{y}) = \epsilon. \tag{7.11}$$

The probability of the event that for at least one i, $X_i\neq Z_i$ is then bounded above by $y\cdot\epsilon$.

Suppose $X\in\Sigma$ is the label of the root of $Ext(T_0)$ in picture 1, and Z that in picture 2, determined by using the recursion function Γ repeatedly upward starting at level s. Then $X\sim\Psi_c^s(T_0,\vec{x}),\ Z\sim\Psi_c^s(T_0,\vec{y})$, from Definition 7.3.

Recall, from Definition 7.1, that $BAD[T_0]$ is the event that $Ext(T_0)$ is not s_0-universal. If $GOOD[T_0] = BAD[T_0]^c$, then under $GOOD[T_0]$, the Ehrenfeucht value of $Ext(T_0)$ is completely determined by $Ext(T_0)|_{s_0}$, which is T_0 itself. And as T_0 is fixed, this means that $EV[Ext(T_0)]$ is then independent of $\vec{x}, \vec{y}$. Thus

$$\begin{aligned} TV(\Psi_c^s(T_0,\vec{x}), \Psi_c^s(T_0,\vec{y})) \leq & P[X \neq Z] \\ \leq & \sum_{y=0}^{\infty} P[Y=y] \cdot y \cdot \epsilon \cdot \mathbf{1}_{BAD[T_0]} \\ = & E[Y \ \mathbf{1}_{BAD[T_0]}]\epsilon. \end{aligned}$$

From Lemma 7.2 and Lemma 7.6,

$$TV(\Psi_c^s(T_0,\vec{x}), \Psi_c^s(T_0,\vec{y})) \leq A(t)\epsilon \text{ with } A(t) = kte^{-t\kappa}. \tag{7.12}$$

Here $A = A(t)$ approaches zero as $t \to \infty$ and so there exists K_0 such that $A \leq K_0$ for *any* choice of t. $\square$

7.5 Proving contraction

We first show Theorem 3.2 in terms of the TV metric. Pick s_0 sufficiently large so that, say, the probability that $T|_{s_0}$ is not s_0-universal is at most $(2K_0)^{-1}$, K_0 given by Lemma 7.8. This can be done because of Lemma 5.7. Let $\vec{x}, \vec{y} \in D$ with $\epsilon = TV(\vec{x}, \vec{y})$. We bound $TV(\Psi_c^s(\vec{x}), \Psi_c^s(\vec{y}))$ by Lemma 7.4. Consider $TV(\Psi_c^s(T_0,\vec{x}), \Psi_c^s(T_0,\vec{y}))$. When T_0 is s_0-universal this has value zero. Otherwise its value is bounded by $K_0\epsilon$ by Lemma 7.8. Lemma 7.4 then gives

$$TV(\Psi_c^s(\vec{x}), \Psi_c^s(\vec{y})) \leq \frac{1}{2K_0} K_0 \epsilon \leq \frac{\epsilon}{2}. \tag{7.13}$$

Finally, we switch to the L^2 metric. For B a sufficiently large constant the inequalities (3.7)978-1-4419-3157-3 yield, say,

$$||\Psi_c^{sB}(\vec{x}) - \Psi_c^{sB}(\vec{y})||_2 \leq \frac{1}{2}||\vec{x} - \vec{y}||_2. \tag{7.14}$$

Then Theorem 3.2 is satisfied with s replaced by sB and $\alpha = \frac{1}{2}$.

8 Implicit function

Here we deduce Theorem 2.4 and hence Theorem 1.2, that $\Pr[A]$ is always a C^∞ function of c. This follows from three results:

1. The function $\Delta(c, \vec{x}) = \Psi_c(\vec{x})$ has all derivatives of all orders.
2. For each $c > 0$ the function
$$F(\vec{x}) = \Psi_c(\vec{x}) - \vec{x} \tag{8.1}$$
has a unique zero $\vec{x} = \vec{x}(c)$.
3. The function $\Psi_c : D \to D$ is contracting in the sense of Theorem 3.2.

(1) is discussed in Subsection 2.4, where we first define $\Psi_c(\vec{x})$, whereas (2) is addressed in Theorem 6.1. The proof of (3) can be found in two parts: the subcritical case in Subsection 3.3, and the supercritical case in Section 7.

Let A be the Jacobian of Ψ_c at $\vec{x}(c)$. From Property 3 all of the eigenvalues of A lie inside the complex unit circle. Then $A - I$ is the Jacobian of F from Property 2. Then $(A - I)^{-1} = -\sum_{u=0}^{\infty} A^u$ is a convergent sequence, and so $A - I$ is invertible. As by Property 1 the function Δ is smooth, the Implicit Function Theorem gives that the fixed point function $\vec{x}(c)$ of F is C^∞.

References

[1] N. Immerman: Descriptive complexity, *Springer Science & Business Media*, 2012; ISBN-13: 978-0387986005. MR-1732784

[2] D. A. Levin, Y. Peres and E. L. Wilmer: Markov Chains and Mixing Times, *American Mathematical Society*, 2006; ISBN: 978-0-8218-4739-8. MR-2466937

[3] D. Marker: Model theory: an introduction, *Springer Science & Business Media*, 2006; ISBN: 978-1-4419-3157-3.

[4] M. Podder and J. Spencer: First Order Probabilities For Galton-Watson Trees. To appear in A Journey Through Discrete Mathematics: A Tribute to Jiří Matoušek; *Springer International Publishing*, 2017; ISBN 978-3-319-44478-9.

[5] J. Spencer: The Strange Logic of Random Graphs; *Springer Publishing Company, Inc.*, 2010; Series: Algorithms and Combinatorics, Vol. 22; ISBN:3642074995 9783642074998. MR-1847951