On the Very Weak 0-1 Law for Random Graphs with Orders

Saharon Shelah

The Hebrew University, Math Institute

Rutgers University, Math Department

Abstract: Let us draw a graph R on $\{0, 1, \dots, n-1\}$ by having an edge $\{i, j\}$ with probability $p_{|i-j|}$, where $\sum_i p_i < \infty$, and let $M_n = (n, <, R)$. For a first order sentence ψ let a^n_ψ be the probability of $M_n \models \psi$. We know that the sequence $a^1_\psi, a^2_\psi, \dots, a^n_\psi, \dots$ does not necessarily converge. But here we find a weaker substitute which we call the very weak 0-1 law. We prove that $\lim_{n\to\infty}(a^n_\psi - a^{n+1}_\psi) = 0$. For this we need a theorem on the (first order) theory of distorted sum of models.

Saharon: 1) Check: line before rest of the proof 3.4A, should it be $\neq$?

2) Check the numerical bounds.

Research partially supported by the Binational Science Foundation and partially sponsored by the Edmund Landau Center for research in Mathematical Analysis, supported by the Minerva Foundation (Germany). Done fall 91, Publication 463.

§0 Introduction

The kind of random models $M_n = (n, <, R)$ from the abstract are from Luczak Shelah [LuSh 435] where among other things, it is proved that the probability $a^n_\psi =: \mathrm{Prob}(M_n \models \psi)$ of $M_n \models \psi$ does not necessarily converge (but if $\sum_i ip_i < \infty$ then it converges, and the value of $\sum p_i$ is not enough, in general, to determine convergence). The theorem in the abstract appears in §1 and is proved in §3, it says that the sequence of probabilities still behaves (somewhat) nicely.

The first results (in various probabilistic distributions) on the asymptomatic behavior of a^n_ψ (see Glebski et al [GKLT], Fagin [F] and survey [Sp]) say that it converges to 0 or 1 hence the name zero one law. In other cases weaker results were gotten: a^n_ψ converges (to some real). We suggest an even weaker version: $|a^{n+1}_\psi - a^n_\psi|$ converges to zero. We also define h-very weak zero one law (see Definition 1.2(2)), but concentrate on the one above. Note that many examples to nonconvergence are done by finding ψ such that e.g. if $\log(\mathrm{n}) = 1 \bmod 10$ then $a^n_\psi \sim 1$ and if $\log(\mathrm{n}) = 5 \bmod 10$ then $a^n_\psi \sim 0$ (or even using functions with $h(n) \to \infty$, $h(n) << \log(\mathrm{n})$). As the most known results were called zero one law we prefer the name "very weak 0-1 law" on weak convergence. A (first order) sentence whose probability a^n_ψ (defined above, for the distribution defined above) may not converge (by [LuSh 435]) is $\psi_0 =: (\exists x)\forall yz(y < x \le z \to \neg yRz)$. But we can find a sequence $m_0 < m_1 < \ldots$ such that the probability that $\Phi =: \bigvee_i (\exists yz)(y \le m_i \,\&\, m_{i+1} \le z \,\&\, yRz)$ is very small. How do we prove $0 = \lim(a^n_\psi - a^{n+1}_\psi)$? By changing the rule of making the random choice we get M'_n nice enough, ensuring Φ holds, while the probability changes little (see §3). Now M'_n is almost the sum of $M'_n \restriction (m_i, m_{i+1})$, precisely M'_n is determined by $M'_n \restriction [m_i, m_{i+2})$, for $i = 0, 1, 2, \ldots$, so we call it a distorted sum. Now a model theoretic lemma from §2 on the n-theory of a distorted sum of models enables us to prove the main theorem in §3 (on the model theoretic background see §2). Later in §4, §5 we deal with some refinements not needed for the main theorem (1.4).

In §1 we also get the very weak 0-1 law for a random partial order suggested by Luczak. In a subsequent paper [Sh 548] we prove the very weak zero law for some other very natural cases: e.g. for a random 2-place function and for $(n, <, R)$ with $<$ the natural order, R a random graph (=symmetric irreflexive relation) with edge probability p. In another one, [Sh 467] we deal with zero one law for the random model from the abstract with $p_i = \frac{1}{i^a}$ (mainly: no order, $a \in (0, 1)$ irrational). See also [Sh 550], [Sh 551], [Sh 581]. Spencer is continuing [Sh 548] looking at the exact h for which h-very weak zero one law holds (see Definition 1.2(2) here). I thank Shmuel Lifsches for many corrections.

Notation

$\mathbb{N}$ is the set of natural numbers.
We identify $n \in \mathbb{N}$ with the set $\{0, \ldots, n-1\}$.
$\mathbb{Z}$ is the set of integers.
$\mathbb{R}$ is the set of reals, $\mathbb{R}^+$ is the set of reals which are positive (i.e. > 0).
$i, j, k, \ell, m, n, r, s, t$ are natural numbers.
ε, ζ are positive reals (or functions with values in $\mathbb{R}^+$).
f, g, h are functions.
τ denotes a vocabulary (for simplicity- set of predicates), $\mathcal{L}^{\rm fo}_\tau$ is the set of first order formulas in the vocabulary τ. (Generally, if $\mathcal{L}$ is a logic $\mathcal{L}_\tau$ the set of sentences (or formulas) in $\mathcal{L}$ in the vocabulary τ and $\mathcal{L}^{\rm fo}$ is the first order logic).
For a first order sentence (or formula) φ let $d_\varphi = d[\varphi]$ be its quantifier depth.
M, N denote models, but we do not distinguish strictly between a model and its universe = set of elements.
$\tau(M)$ is the vocabulary of M, for $R \in \tau(M)$, $n(R)$ is the number of places (=arity of R), R^M the interpretation of R in M.
A basic formula is one of the form $\pm R(x_{i_0}, \ldots, x_{i_{n(R)-1}})$ (i.e. $R(x_{i_0}, \ldots, x_{i_{n(R)-1}})$ or $\neg R(x_{i_0}, \ldots, x_{i_{n(R)-1}})$
$\bar a$ denotes a sequence of elements of a model. $\ell g(\bar a)$ is the length of $\bar a$.

If $<$ belongs to the vocabulary τ, then in τ-models $<^M$ is a linear order, if not said otherwise. If M_i are τ-models for $i < n$, $M = \sum_{i<n} M_i$ is (assuming for simplicity the universes are pairwise disjoint) the models defined by: universe $\bigcup_{i<n} M_i$, $R^M = \bigcup_{i<n} R^{M_i}$ for $R \in \tau$ except that if $< \in \tau$ then: $x <^M y \Leftrightarrow \bigvee_{i<j<n} [x \in M_i \,\&\, y \in M_j] \vee \bigvee_{i<n} x <^{M_i} y$ (similarly with any linear order I instead of n). We write $M_0 + M_1$ instead of $\sum_{i<2} M_i$.
$\psi^{\text{if}(\theta)}$ is ψ if θ is true, $\neg\psi$ if θ is false.
We identify true, false with yes, no.
Note: t is a natural number, $\mathbf{t}$ kind of depth n theory, $\mathfrak{t}$ is a truth value, $\blacktriangle$ is the symmetric difference, Δ denote a set of formulas.

§1 The Very Weak Zero One Law.

1.1 Definition: 1) A 0-1 law context is a sequence $\bar{K} = \langle K_n, \mu_n : n < \omega \rangle$ such that:

(a) for some vocabulary $\tau = \tau_{\bar{K}}$, for every n, K_n is a family of τ-models, closed under isomorphism with the family of isomorphism types being a set.

(b) for each n, μ_n is a probability measure on the set of isomorphism types (of models from K_n).

2) $\bar{K}$ is finitary (countable) if for each n the set $\{M/\cong : M \in K_n\}$ is finite (countable).

3) For a sentence ψ (not necessarily f.o.) $\mathrm{Prob}_{\mu_n}(M_n \models \psi)$ or $\mathrm{Prob}_{K_n}(M_n \models \psi)$ or $\mathrm{Prob}_{\mu_n}(M_n \models \psi \Big| M_n \in K_n)$ means $\mu_n\{M/\cong : M \in K_n, M \models \psi\}$.

4) Instead of clause (a) of (1), we may use K_n a set of τ-models, μ_n a probability measure on K_n; particularly we introduce a random choice of M_n; the translation between the two contexts should be clear.

1.1A Discussion: A 0-1 law context is not necessarily a context which satisfies a 0-1 law. It is a context in which we can formulate a 0-1 law, and also weaker variants.

1.2 Definition: 0) A 0-1 context $\bar{K}$ satisfies the 0-1 law for a logic $\mathcal{L}$ if for every sentence $\varphi \in \mathcal{L}_\tau$ (with $\tau = \tau_{\bar{K}}$ of course) we have: $a^n_\varphi \stackrel{\mathrm{def}}{=} \mathrm{Prob}_{\mu_n}(M_n \models \varphi)$ converges to zero or converges to 1 when $n \to \infty$.

1) $\bar{K}$ satisfies the very weak 0–1 law for the logic $\mathcal{L}$ if for every sentence $\varphi \in \mathcal{L}_\tau$ we have: $a_n \stackrel{\mathrm{def}}{=} \mathrm{Prob}_{\mu_{n+1}}(M_{n+1} \models \varphi \Big| M_{n+1} \in K_{n+1}) - \mathrm{Prob}_{\mu_n}(M_n \models \varphi \Big| M_n \in K_n)$ converges to zero as $n \to \infty$.

2) $\bar{K}$ satisfies the h-very weak 0–1 one law for the logic $\mathcal{L}$ if for every sentence $\varphi \in \mathcal{L}_\tau$, $\max_{m_1,m_2 \in [n,h(n)]} \Big[\mathrm{Prob}_{\mu_{m_1}}(M_{m_1} \models \varphi \Big| M_{m_1} \in K_{m_1}) - \mathrm{Prob}_{\mu_{m_2}}(M_{m_2} \models \varphi \Big| M_{m_2} \in K_{m_2})\Big]$ converge to zero as $n \to \infty$. (We shall concentrate on part 1).

3) $\bar{K}$ satisfies the convergence law for $\mathcal{L}$ if for every $\varphi \in \mathcal{L}_\tau$ we have: $\langle \mathrm{Prob}_{\mu_n}(M_n \models \varphi \Big| M_n \in K_n) : n < \omega \rangle$ converges (to some real $\in [0,1]_{\mathbb{R}}$). (So if it always converges to 0 or to 1 we say that $\bar{K}$ satisfies the 0-1 law for $\mathcal{L}$).

4) If $\mathcal{L}$ is the first order logic we may omit it.

The following $\bar{K}$ is from Luczak Shelah [LuSh 435].

1.3 Definition: Let $\bar{p} = \langle p_i : i \in \mathbb{N} \rangle$, p_i a real number $0 \leq p_i \leq 1$, $p_0 = 0$. We define $\bar{K}^{\mathrm{og}}_{\bar{p}}$ as follows:
the models from K_n are of the form $M = (n, <, R)$, so we are using the variant

from[†] Definition 1.1(4) where $n = \{0, \ldots, n-1\}$, $<$ the usual order, R a graph (i.e. a symmetric relation on n which is irreflexive i.e. $\neg xRx$) and: $\mathrm{Prob}_{\mu_n}(M_n \cong M \big| M_n \in K_n)$ is $\prod\{p_{j-i} : i < j < n \text{ and } iRj\} \times \prod\{1 - p_{j-i} : i < j < n \text{ and } \neg iRj\}$, i.e. for each $i < j < n$ we decide whether iRj by flipping a coin with probability p_{j-i} for yes, independently for the different pairs.

1.4 Theorem: $\bar{K}^{\mathrm{og}}$ satisfies the very weak 0–1 law <u>if</u> $\sum_i p_i < \infty$.

This is our main result. We shall prove it later in §3. Luczak [Lu] suggested another context:

1.5 Definition: $\bar{K}^{\mathrm{opo}}_{p_n}$ is the following 0-1 law context (p_n is a function of n, $0 \le p_n \le 1$; so more precisely we are defining $\langle K^{\mathrm{opo}}_{p_n} : n < \omega \rangle$). The models are of the form, $(n, <, <^*)$, $n = \{0, 1, \ldots, n-1\}$, $<$-the usual order, $<^*$-the following partial order: we draw a random graph on n (edge relation R) with edge probability p_n, $x <^* y$ iff there are $k \in \mathbb{N}$ and $x = x_0 < x_1 < \ldots < x_k = y < n$ such that $x_\ell R x_{\ell+1}$. The probability is derived from the probability distribution for R (which does not appear in the model.)

1.6 Theorem Assume $p_n = \frac{1}{(n+1)^a}$, where $0 < a < 1$ (like [ShSp 304]). Then $\bar{K}^{opo}_{p_n}$ satisfies the very weak zero one law.

Proof: Similar to the previous theorem as easily
(*) for $\varepsilon > 0$ we have: $\mathrm{Prob}(M_n \models (\exists x < y)[\neg x <^* y \& x + n^{1-\varepsilon} \le y)$ is very small. $\square_{1.6}$

§2 Model Theory: Distorted sum of Models

The main lemma (2.14) generalizes the addition theory and deals with models with distances (both from model theory). Concerning the first, see Feferman Vaught [FV]. The method has its origin in Mostowski [Mo], who dealt with reduced products. The first work on other products is Beth [B] who dealt with the ordinal sum of finitely many ordered systems. For a presentation and history see Feferman Vaught [FV], pp 57–59 and Gurevich [Gu] (particularly the th^n's). Concerning models with distance see Gaifman [Gf], a forerunner of which was Marcus [M1] who deals with the case of $M = (M, F, P_i)_{i<n}$, F a unary function, P_i unary predicates and the distance is as in the graph which the function F

[†] but no two models are isomorphic so the difference is even more trivial

defines (i.e. x, y connected by an edge if: $x = F(y)$ or $y = F(x)$; where for a graph G the distance is $d_G(x, y) = \min\{k$: we can find $x_0, \ldots, x_k$ such that: $x = x_0$, $y = x_k$, and $x_\ell, x_{\ell+1}$ are connected$\}$). We may look at our subject here as dealing with sums with only local "disturbances", "semi sums"; "distorted sums". The connections are explained in §4, §5.

In 2.16 we draw the conclusion for linear order which we shall use in §3 to prove theorem 1.4, in fact proving the main theorem 1.4 from §2 we use almost only 2.16. Elsewhere we shall return to improving the numerical bound involved in the proof see [Sh, F-120]

Note: $\mathfrak{B}$ is a (two sorted) model.

2.1 Definition: 1) We call σ a vocabulary of systems if $\sigma = \langle \tau_1, \tau_2 \rangle$, τ_1, τ_2 are sets of predicates (usually finite but not needed).

2) We call $\mathfrak{B}$ a σ-system if:

(A) $\mathfrak{B} = (M, I, h, d)$ $(= (M^{\mathfrak{B}}, I^{\mathfrak{B}}, h^{\mathfrak{B}}, d^{\mathfrak{B}}))$.

(B) M is a τ_1-model and I is a τ_2-model, but we use M, I also for their universes.

(C) h is a function from M onto I,

(D) d is a distance function on I, i.e.

- (α) d is a symmetric two place function from I to $\mathbb{N} \cup \{\infty\} = \{0, 1, 2, 3, \ldots, \infty\}$,
- (β) $d(x, x) = 0$
- (γ) $d(x, z) \leq d(x, y) + d(y, z)$.

(E) M, I are disjoint.

3) Let $\sigma(\mathfrak{B}) = \sigma$ for $\mathfrak{B}$ a σ-system and $\tau_\ell(\mathfrak{B}) = \tau_\ell$ if $\sigma = \langle \tau_1, \tau_2 \rangle$.

2.1A Discussion: The demands "M, I are disjoint and h is onto I" are not essential, this is just for convenience of presentation. If h is not onto I, we should allow relations on $M \cup I$, but then if M, I are not disjoint then for each predicate, for each of its places we should assign: does there appear a member of M or a member of I; also we should have two kinds of variables not allowing equality between variables of the different kinds. So in application we may use those alternative presentations.

2.2 Conventions: (1) We may allow function symbols (and individual constants) but then treat them as relations.

(2) We stipulate $h(x) = x$ for $x \in I$ (remember 2.1(2)(E)).

(3) $\bar{a} \subseteq \mathfrak{B}$ means $\mathrm{Rang}\,(\bar{a}) \subseteq M^{\mathfrak{B}} \cup I^{\mathfrak{B}}$.

(4) A model M will be identified with the system $\mathfrak{B} = \mathfrak{B}^{\text{sim}}[M] : M^{\mathfrak{B}} = M$, $I^{\mathfrak{B}} = M$, $h^{\mathfrak{B}} = \text{id}_M$,

$$d(x, y) = \begin{cases} 0 & x = y \\ \infty & x \neq y \end{cases}$$

(you may of course take two disjoint copies of M as $M^{\mathfrak{B}}$ and $I^{\mathfrak{B}}$).
(5) From a model M we can derive another system $\mathfrak{B}^{\text{dis}}[M]$: $M^{\mathfrak{B}} = M$,$I^{\mathfrak{B}} = M$, $h^{\mathfrak{B}} = \text{id}_M$, and $d(x, y) \overset{\text{def}}{=} \text{Min}\{n$: there are $z_0, \ldots, z_n \in M$, $x = z_0, y = z_n$ and for $\ell < n$ for some $R \in \tau(M)$, and sequence $\bar{a} \in R^M$ we have $\{z_\ell, z_{\ell+1}\} \subseteq \text{Rang}\, \bar{a}\}$ (remember $\tau(M)$ is the vocabulary of M; this is the definition of distance in Gaifman [Gf]).

2.2A Remark: 1) So the difference between $\mathfrak{B}^{\text{sim}}[M]$ and $\mathfrak{B}^{\text{dis}}[M]$ is only in the choice of the distance function d.
2) Below in $\mathcal{L}_{\sigma_0}$ the formula $d(x, y) = 0$ appears which normally means $x = y$. If not we could have in Def. 2.3(3), (4) replaced "$k \leq r$", "$s \leq n$" by "$k < r$", "$s < n$" respectively.

2.3 Definition: 1) For a system $\mathfrak{B} = (M, I, h, d)$:
for $x \in I$,

$N_r(x) \overset{\text{def}}{=} \{y \in I : d(y, x) \leq r\}$

for $x \in M \cup I$,

$N_r^+(x) \overset{\text{def}}{=} \{y \in M \cup I : d(h(y), h(x)) \leq r\}$.

2) $\mathcal{L}_\sigma$ is the set of first order formulas for σ; we have variables on M, variables on I, the predicates of τ_1, τ_2; and the additional atomic formulas $h(x) = y$; and "$h(y) \in N_r(h(x))$" for each r (see part (4) below).
3) $\mathfrak{B}^- = (M^{\mathfrak{B}}, I^{\mathfrak{B}}, h)$, $\mathfrak{B}_r = (M^{\mathfrak{B}}, I^{\mathfrak{B}}, h^{\mathfrak{B}}, \text{“}d(x, y) \leq k\text{”})_{k \leq r}$ so $\mathfrak{B}_0 = \mathfrak{B}^-$ and we let $\sigma_r = \sigma_r(\mathfrak{B}) = \sigma(\mathfrak{B}_r)$ (so, $\mathfrak{B}_r$ is a two sorted model but not a system).
4) So $\mathcal{L}_{\sigma_n}$ is defined as in part (2) but "$h(y) \in N_s(h(x))$" appear only for $s \leq n$.

* * *

We usually apply our theorems in the following case:

2.4 Definition: A system $\mathfrak{B}$ is *simple* if
(a) $d(x, y) \leq 1$ (where x, y vary on I) is equivalent to a quantifier free formula in $\mathcal{L}_{\sigma_0(\mathfrak{B})}$.
(b) for $x, y \in I$ we have: $d(x, y) \leq r$ *iff* there are $x_0, \ldots, x_r \in I$ such that $x = x_0$, $y = x_r$, and $d(x_\ell, x_{\ell+1}) \leq 1$ for $\ell < r$ (i.e., like 2.2(5)).

2.4A Remark: If $\mathfrak{B}$ is simple, note that every formula in $\mathcal{L}_{\sigma(\mathfrak{B})}$ is equivalent to some formula in $\mathcal{L}_{\sigma_0}$, if we know just that clause (b) of Definition 2.4 is satisfied then every formula in $\mathcal{L}_{\sigma(\mathfrak{B})}$ is equivalent to some formula in $\mathcal{L}_{\sigma_1}$.

2.5 Convention: 1) We define f: we let $f_n(r) = r + 3^n$ for $r, n \in \mathbb{N}$ <u>or</u> more generally, f a two place function (written $f_n(r)$) from $\mathbb{N}$ to $\mathbb{N}$ satisfying: f_n non decreasing in n and in r, $r < f_n(r) \in \mathbb{N}$ and $f_n^{(3)}(r) \leq f_{n+1}(r)$ where $f_n^{(0)}(r) = r$, $f_n^{(\ell+1)}(r) = f_n(f_n^{(\ell)}(r))$ and $f_n^{(2)}(r) \geq f_n(r) + f_n(0)$.
2) We call f *nice* if in addition $f_n^{(4)}(r) \leq f_{n+1}(r)$.
3) For g a function from $\mathbb{N}$ to $\mathbb{N}$, let $g(\langle r_\ell : \ell < m \rangle) = \langle g(r_\ell) : \ell < m \rangle$.

2.6 Definition: 1) For a system $\mathfrak{B} = (M, I, h, d)$ and $m, n \in \mathbb{N}$, and $\bar{a} = \langle a_0, \dots, a_{m-1} \rangle \subseteq \mathfrak{B}$ we define $\text{th}_r^n(\bar{a}, \mathfrak{B})$, here r stands for a distance. We define it by induction† on n:

$$\text{th}_r^0(\bar{a}, \mathfrak{B}) \text{ is } \{\varphi(x_0, \dots, x_{m-1}) : (M, I, h) \models \varphi[a_0, \dots, a_{m-1}], \ \varphi \text{ a basic formula in } \mathcal{L}_{\sigma(\mathfrak{B}_r)}\}$$
$$\text{th}_r^{n+1}(\bar{a}, \mathfrak{B}) = \{\text{th}_r^n(\bar{a}\frown\langle c \rangle, \mathfrak{B}) : c \in M \cup I\}.$$

2) If σ is a vocabulary of systems, $n, m \in \mathbb{N}$, <u>then</u> $\text{TH}_r^n(m, \sigma)$ is the set of formally possible $\text{th}_r^n(\bar{a}, \mathfrak{B})$ for $\mathfrak{B}$ a σ-system, $\bar{a} \subseteq \mathfrak{B}$ and $\lg(\bar{a}) = m$; this is defined naturally. Pedantically, we define it by induction on n; <u>for $n = 0$</u> it is the family of sets $\mathbf{t}$ of basic formulas $\varphi(x_0, \dots, x_{m-1})$ of $\mathcal{L}_{\sigma_r}$ such that for each atomic $\varphi(x_0, \dots, x_{m-1})$ exactly one of $\varphi(x_0, \dots, x_{m-1}), \neg\varphi(x_0, \dots, x_{m-1})$ belongs to $\mathbf{t}$. $\text{TH}_r^{n+1}(m, \sigma)$ is the family of subsets of $\text{TH}_r^n(m+1, \sigma)$.
3) If τ is a vocabulary of models (see 2.2(4)), $n, m \in \mathbb{N}$ <u>then</u> $\text{TH}^n(m, \tau)$ is the set of formally possible $\text{th}_0^n(\bar{a}, \mathfrak{B})$, $\mathfrak{B}$ a τ-model, i.e. $\mathfrak{B} = \mathfrak{B}^{\text{sim}}[M]$ for some τ-model M (note: the value of r is immaterial as the distance function is trivial).
4) If $r = 0$ we may omit it, so for a model M, using $I = M$, h the identity we get the usual $th^n(\bar{a}, M)$ (but we do not assume knowledge about it).
5) If $\bar{a}$ is empty sequence we may omit it.

2.7 Claim: 1) For $\mathfrak{B}, n, m$, $\bar{a}$ as above, $\varphi = \varphi(x_0, \dots, x_{m-1})$ a (first order) formula in $\mathcal{L}_{\sigma(\mathfrak{B}_r)}$ of quantifier depth n, we have:
from $\text{th}_r^n(\bar{a}, \mathfrak{B})$ we can compute the truth value of "$\mathfrak{B}_r \models \varphi[\bar{a}]$".
Here and in later instance we mean:
for any $\mathbf{t} \in \text{TH}_r^n(m, \sigma)$ we can compute a truth value $\mathbf{t}$ such that: if $\mathbf{t} = \text{th}_r^n(\bar{a}, \mathfrak{B})$

† In the following definition basic is atomic or a negation of atomic.

then $\mathbf{t}$ is the truth value of "$\mathfrak{B}_r \models \varphi[\bar{a}]$". Also in the proof we behave similarly.
2) For any σ and $r, n, m \in \mathbb{N}$, if $\mathbf{t} \in \mathrm{TH}^n_r(m, \sigma)$ <u>then</u> for some formula $\varphi(x_0, \dots, x_{m-1}) \in \mathcal{L}_{\sigma_r}$ of quantifier depth n, for any σ-system $\mathfrak{B}$ and $\bar{a} = \langle a_0, \dots, a_{m-1}\rangle \subseteq \mathfrak{B}$ we have: $\mathbf{t} = \mathrm{th}^n_r(\bar{a}, \mathfrak{B})$ iff $\mathfrak{B} \models \varphi[\bar{a}]$.
3) The functions $\langle \sigma, n, m, r\rangle \mapsto \mathrm{TH}^n_r(m, \sigma)$ and $\langle \tau, n, m\rangle \mapsto \mathrm{TH}^n(m, \tau)$ are computable.
4) From $\mathrm{th}^n_r(\bar{a}, \mathfrak{B})$ we can compute $\mathrm{th}^m_s(\bar{a}, \mathfrak{B})$ if $r \geq s$ and $n \geq m$. Also if $\mathrm{Rang}\,(\bar{b}) \subseteq \mathrm{Rang}\,(\bar{a})$ then from $\mathrm{th}^n_r(\bar{a}, \mathfrak{B})$ and $\{(\ell, m) : b_\ell = a_m\}$ we can compute $\mathrm{th}^n_r(\bar{b}, \mathfrak{B})$. (See part (1).)
5) If $\mathfrak{B}$ is simple <u>then</u> from $\mathrm{th}^{n+r}_1(\bar{a}, \mathfrak{B})$ we can compute $\mathrm{th}^n_{2r}(\bar{a}, \mathfrak{B})$.
6) If $n_1, n_2 \geq 2^d$ <u>then</u> $\mathrm{th}^d(\langle\rangle, (n_1, <)) = \mathrm{th}^d(\langle\rangle, (n_2, <))$. Also if $\mathrm{th}^d(\langle\rangle, M^\ell_i) = \mathbf{t}$ for $\ell = 1, 2$ and for $i < \max\{n_1, n_2\}$, τ a vocabulary, $M_i \in K_\tau$ *then* $\mathrm{th}^d(\langle\rangle, \sum_{i<n_1} M^1_i) = \mathrm{th}^d(\langle\rangle, \sum_{i<n_2} M^2_i)$. If $M^\ell_i \in K_\tau$ for $\ell = 1, 2$, $i < k$ and $\mathrm{th}^d(M^1_i) = \mathrm{th}^d(M^2_i)$ for $i < k$ then $\mathrm{th}^d(\langle\rangle, \sum_{i<k} M^1_i) = \mathrm{th}^d(\langle\rangle, \sum_{i<k} M^2_i)$.
7) For a given vocabulary τ, and $d \in \mathbb{N}$ there is an operation $\oplus$ on $\mathrm{TH}^d(0, \tau)$ such that $\mathrm{th}^d(\langle\rangle, \sum_{i<k} M_i) = \oplus\langle \mathrm{th}^d(\langle\rangle, M_i) : i < k\rangle$, this operation is associative (but in general not commutative).

Proof: 1) We prove this by induction on the formula. (It goes without saying that the reasoning below does not depend on $\mathfrak{B}$.)

φ *atomic:* Thus $n = d(\varphi) = 0$, and by the Definition of $\mathrm{th}^n_r(\bar{a}, \mathfrak{B})$ the statement is trivial.

$\varphi = \neg\psi$*:* Easy by the induction hypothesis.

$\varphi = \varphi_1 \wedge \varphi_2$ (or $\varphi_1 \vee \varphi_2$, or $\varphi_1 \rightarrow \varphi_2$): Easy by the induction hypothesis.

$\varphi = (\exists x)\varphi_1$*:* Without loss of generality $\varphi = (\exists x_m)\varphi_1(x_0, \dots, x_{m-1}, x_m)$. So $d(\varphi_1) = n - 1$, and by the induction hypothesis for $a_0, \dots, a_m \in \mathfrak{B}$ we have: the truth value of $\mathfrak{B} \models \varphi_1[a_0, \dots, a_{m-1}, a_m]$ is computable from $\mathrm{th}^{n-1}_r(\langle a_0, \dots, a_m\rangle, \mathfrak{B})$. Say it holds iff $\mathrm{th}^{n-1}_r(\langle a_0, \dots, a_m\rangle, \mathfrak{B}) \in \mathbf{T}_{\varphi_1}$ ($\mathbf{T}_{\varphi_1}$ a subset of $\mathrm{TH}^{n-1}_r(m + 1, \sigma(\mathfrak{B}))$).
Now $\mathfrak{B} \models \varphi[a_0, \dots, a_{m-1}]$ iff for some a_m, $\mathfrak{B} \models \varphi_1[a_0, \dots, a_{m-1}, a_m]$, iff $\mathrm{th}^{n-1}_r(\langle a_0, \dots, a_{m-1}, a_m\rangle, \mathfrak{B}) \in \mathbf{T}_{\varphi_1}$ ($\mathbf{T}_{\varphi_1}$ the subset of $\mathrm{TH}^{n-1}_r(m+1, \sigma(\mathfrak{B}))$ from above) for some $a_m \in \mathfrak{B}$ *iff* for some $c \in \mathfrak{B}$, $\mathrm{th}^{n-1}_r(\langle a_0, \dots, a_{m-1}\rangle^\frown\langle c\rangle, \mathfrak{B}) \in \mathbf{T}_{\varphi_1}$ *iff* $\mathrm{th}^n_r(\langle a_0, \dots, a_{m-1}\rangle, \mathfrak{B})$ is not disjoint to $\mathbf{T}_{\varphi_1}$ (the first "iff" by the definition of satisfaction, the second "iff" by the choice of $\mathbf{T}_{\varphi_1}$, the third "iff" is trivial; the last "iff" by the induction step in the definition of th^n_r). So we have completed

the induction.
2) We define $\varphi = \varphi_{\mathbf{t}}$ for $\mathbf{t} \in \mathrm{TH}^n_r(m,\sigma)$ as required by induction on n; check the inductive definition of th^n_r.
3) Read the definition. (2.6(2))
4) By induction on n, for $n = 0$ as th^n_r "speaks" on more basic formulas. For $n+1$ using the induction hypothesis (and the definition of th^{n+1}_r).
5) We prove it by induction on n. The step from n to $n+1$ is very straightforward. For $n = 0$, we prove the statement by induction on r. For $r = 0$ note $\mathrm{th}^n_{2^r}(\bar a, \mathfrak{B}) = \mathrm{th}^0_{2^0}(\bar a, \mathfrak{B}) = \mathrm{th}^0_1(\bar a, \mathfrak{B})$ so there is nothing to prove. For $r = r(0)+1$ just note that for $s_0 \le 2^r$ we have: $d(x,y) \le s_0$ is equivalent to $\bigvee_{\substack{s_1+s_2=s_0\\ s_1,s_2\le 2^{r(0)}}} (\exists z)[d(x,z) \le s_1 \,\&\, d(z,y) \le s_2]$
6) The first phrase is a special case of the second (with M_i a model with a single element: i). Let $n(\ell) \stackrel{\mathrm{def}}{=} n_\ell$. For the second phrase we prove the following more general statement by induction on d:
$(*)_d$ Assume that for $\ell = 1,2$ we have:

$$M_\ell = \sum_{i<n(\ell)} M^\ell_i \text{ and } 0 \le i_\ell(1) < i_\ell(2) < \ldots < i_\ell(k^*-1) < n(\ell),$$

we stipulate $i_\ell(0) = -1$ and $i_\ell(k^*) = n(\ell)$ (possibly $k^* = 1$), assume further $\bar a^\ell_k \subseteq M^\ell_{i_\ell(k)}$ has length $m(k)$ for $k < k^*$. Also assume that for each $k = 0,\ldots,k^*-1$ we have $\mathrm{th}^d(\bar a^1_k, M^1_{i_1(k)}) = \mathrm{th}^d(\bar a^2_k, M^2_{i_2(k)})$ and $(i_1(k+1) - i_1(k) - 1)$ and $(i_2(k+1) - i_2(k) - 1)$ are equal or both $\ge 2^d - 1$. Lastly assume $\mathrm{th}^d(M^1_i) = \mathrm{th}^d(M^2_j)$ at least when $(\exists m)[i \in (i_1(m), i_1(m+1))\& j \in (i_2(m), i_2(m+1)]$ (holds automatically when proving second phrase of (6)). Then

$$\mathrm{th}^d(\bar a^1, \sum_{i<n(1)} M^1_i) = \mathrm{th}(\bar a^2, \sum_{i<n(2)} M^2_i) \text{ where } \bar a^\ell = \bar a^\ell_0{}^\frown \bar a^\ell_1{}^\frown \bar a^\ell_2{}^\frown \ldots {}^\frown \bar a^\ell_{k^*-1}.$$

The proof is straightforward, and for the case $k^* = 1$ we get the desired conclusions. Lastly the third phrase of (6) is also a particular case of $(*)_d$: let $n(1) = n(2) = k$, $k^* = n(\ell)+1$, $i_\ell(m) = m-1$, and $\bar a^\ell_i = \langle\rangle$.
7) The proof is like that of (6), but in $(*)_d$ we add $\bigwedge_{k=1}^{k^*-1} i_1(k) = i_2(k)$.

Remark: If we will want to quantify on sequence of elements (i.e. use $\bar c$ rather than c) this helps. $\square_{2.7}$

2.7A Claim: Let $\mathfrak{B}$ be a system, let $\langle a_i : i < m\rangle$ and $\langle r_i : i < m\rangle$ (where $a_i \in \mathfrak{B}$ and $r_i \in \mathbb{N}$) $f_n : \mathbb{N} \to \mathbb{N}$ for $n \in \mathbb{N}$ (not necessarily as in 2.5) be given then for some $\langle n_i : i \in w\rangle$ where $w \subseteq m$ $(= \{0, \dots, m-1\})$ and function g from $\{0, 1, \dots, m-1\}$ to w which is the identity on w we have: $\sum_{i\in w} n_i \le m - |w|$ and the sets in $\langle N^+_{f_{n_i}(\sum\{r_j : g(j)=i\})}(a_i) : i \in w\rangle$ are pairwise disjoint and $\bar a$ is included in their union provided that

(∗) $2f_{n_1}(r_1) + f_{n_2}(r_2) \le f_{n_1+n_2+1}(r_1 + r_2)$ and f non decreasing in r and in n (considering f a two place function from $\mathbb{N}$ to $\mathbb{N}$).

Proof: We call $\langle w, \bar n, g\rangle$ with $\bar n = \langle n_i : i \in w\rangle$ candidate if it satisfies all the requirements in the conclusion of 2.7A except possibly the "pairwise disjoint". Clearly there is a candidate: $w = \{0, \dots, m-1\}$, $\bigwedge_{i\in w} n_i = 0$, g the identity on w. So there is a candidate $\langle w, \bar n, g\rangle$ with $|w|$ minimal. If the disjointness demand holds, then we are done. So assume $i(1) \ne i(2)$ are in w and there is x belonging to $N^+_{f_{n_{i(1)}}(\sum\{r_j:g(j)=i(1)\})}(a_{i(1)})$ and to $N^+_{f_{n_{i(2)}}(\sum\{r_j:g(j)=i(2)\}}(a_{i(2)})$.

Let $w' = w \setminus \{i(2)\}$, g' be a function with domain $\{0, \dots, m-1\}$ defined by: $g'(j)$ is $g(j)$ if $g(j) \ne i(2)$, and $g'(j)$ is $i(1)$ if $g(j) = i(2)$. Lastly define n'_i for $i \in w'$: $n'_i = n_i$ if $i \ne i(1)$ and $n'_i = n_{i(1)} + n_{i(2)} + 1$ if $i = i(1)$ and let $\bar n' = \langle n'_i : i \in w'\rangle$. Now we shall show below that $(w', \bar m', g')$ is a candidate thus finishing the proof: for this we have to check the two relevant conditions. First

$$\sum_{i\in w'} n'_i = \sum_{\substack{i\in w'\\ i\ne i(1)}} n'_i + n'_{i(1)} = \sum_{\substack{i\in w'\\ i\ne i(1)}} n_i + n_{i(1)} + n_{i(2)} + 1 =$$

$$\sum_{i\in w} n_i + 1 \le m - |w| + 1 = m - (|w| - 1) = m - |w'|.$$

Secondly, why $\bigcup_{i\in w'} N^+_{f_{n'_i}(\sum\{r_j:g'(j)=i\})}(a_i)$ includes $\bar a$? if $j(*) < m$ then for some $i \in w$ we have $a_{j(*)} \in N^+_{f_{n_i}(\sum\{r_j:g(j)=i\})}(a_i)$; if $i \ne i(1), i(2)$ then $i \in w'$, and

$$N^+_{f_{n'_i}(\sum\{r_j:g'(j)=i\})}(a_i) = N^+_{f_{n_i}(\sum\{r_j:g(j)=i\})}(a_i)$$

so we are done; if $i = i(1)$ then $n'_i = n_{i(1)} + n_{i(2)} + 1 \ge n_{i(1)} = n_i$ and

$$\sum\{r_j : g'(j) = i\} \ge \sum\{r_j : g(j) = i\}$$

hence

$$N^+_{f_{n'_i}(\sum\{r_j:g'(j)=i\})}(a_i) \supseteq N^+_{f_{n_i}(\sum\{r_j:g(j)=i\})}(a_i)$$

and we are done.

We are left with the case $i = i(2)$, so by the choice of i we have $d(a_{j(*)}, a_{i(2)}) \leq f_{n_{i(2)}}(\sum\{r_j : g(j) = i(2)\})$ and by the choice of x (and $i(1), i(2)$) above $d(a_{i(2)}, x) \leq f_{n_{i(2)}}(\sum\{r_j : g(j) = i(2)\})$ and $d(x, a_{i(1)}) \leq f_{n_{i(1)}}(\sum\{r_j : g(j) = i(1)\})$. So as d is a metric (i.e. the triangular inequality) $d(a_{j(*)}, a_{i(1)}) \leq 2f_{n_{i(2)}}(\sum\{r_j : g(j) = i(2)\}) + f_{n_{i(1)}}(\sum\{r_j : g(j) = i(1)\})$. Now $\sum\{r_j : g'(j) = i(1)\} = \sum\{r_j : g(j) = i(1)\} + \sum\{r_j : g(j) = i(2)\}$ (by the definition of g') and $n'_{i(1)} = n_{i(1)} + n_{i(2)} + 1$, hence what we need is
$(*)$ $2f_{n^1}(r^1) + f_{n^2}(r^2) \leq f_{n^1+n^2+1}(r^1 + r^2)$
which is assumed. $\square_{2.7A}$

2.7B Remark: 1) We can replace in 2.7A and its proof $\sum\{r_j : g(j) = i\}$ by $\max\{r_j : g(j) = i\}$ and (in $(*)$) $r^1 + r^2$ by $\max\{r^1, r^2\}$.
(2) Concerning $(*)$, letting $n = \max\{n^1, n^2\}$ and $r = \max\{r^1, r^2\}$ it suffices to have
$(*)_1$ $f_n(r)$ is non decreasing in n and in r.
$(*)_2$ $3f_n(r) \leq f_{n+1}(r)$

2.8 Definition: 1) For a system $\mathfrak{B}$, and $r, n, m \in \mathbb{N}$, and $\bar{a} = \langle a_0, \dots, a_{m-1}\rangle \subseteq \mathfrak{B}$ we call $\bar{a}$ a $(\mathfrak{B}, r)$-component <u>if</u> $\bar{a} \subseteq N^+_r(a_0)$. In this case we define $\mathrm{bth}^n_r(\bar{a}, \mathfrak{B})$ (bth is for bounded theory). We do it by induction on n (for all r) (the function f from 2.5 is an implicit parameter.)
(α) $\mathrm{bth}^0_r(\bar{a}, \mathfrak{B}) = \mathrm{th}^0(\bar{a}, \mathfrak{B}_r)$
(β) $\mathrm{bth}^{n+1}_r(\bar{a}, \mathfrak{B})$ is $\langle \mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2\rangle$ where

(i) $\mathbf{t}_0 = \mathrm{bth}^n_r(\bar{a}, \mathfrak{B})$
(ii) $\mathbf{t}_1 = \{\mathrm{bth}^n_r(c, \mathfrak{B}) : c \in N^+_{f^{(2)}_n(r)}(a_0) \setminus N^+_{f_n(r)}(a_0)\}$
(iii) $\mathbf{t}_2 = \{\mathrm{bth}^n_{f^{(2)}_n(r)}(\bar{a}\frown\langle c\rangle, \mathfrak{B}) : c \in N^+_{f_n(r)}(a_0)\}$

2) If σ is a vocabulary of systems and $r, n, m \in \mathbb{N}$ then $\mathrm{BTH}^n_r(m, \sigma)$ is the set of formally possible $\mathrm{bth}^n_r(\bar{a}, \mathfrak{B})$, ($\mathfrak{B}$ a σ-system, $\bar{a} = \langle a_0, \dots, a_{m-1}\rangle \subseteq \mathfrak{B}$ and $\bar{a}$ is a $(\mathfrak{B}, r)$-component).

2.9 Claim: 1) For any σ (vocabulary of systems), numbers $n, m, r \in \mathbb{N}$, and $\mathbf{t} \in \mathrm{BTH}^n_r(m, \sigma)$, <u>there is</u> $\varphi = \varphi(x_0, \dots, x_{m-1}) \in \mathcal{L}_\sigma$ (even $\mathcal{L}_{\sigma(\mathfrak{B}_{f_n(r)})}$) of quantifier depth n such that for any σ-system $\mathfrak{B}$, and $\bar{a} = \langle a_0, \dots, a_{m-1}\rangle \subseteq \mathfrak{B}$ a $(\mathfrak{B}, r)$-component we have:

$$\mathbf{t} = \mathrm{bth}^n_r(\bar{a}, \mathfrak{B}) \text{ iff } \mathfrak{B}_{f_n(r)} \restriction N^+_{f_n(r)}(a_0) \models \varphi[\bar{a}].$$

2) If $n \geq m$, $\bar{b}$ is the permutation of $\bar{a}$ by the function h or $\bar{b} = \langle a_\ell : \ell < k \rangle$ for some $k \leq \ell g(\bar{a})$ and $b_0 = a_0$ <u>then</u> from $\mathbf{t} = \text{bth}^n_r(\bar{a}, \mathfrak{B})$ we can compute $\text{bth}^m_r(\bar{b}, \mathfrak{B})$ (using n, m, r, and h or k).

Proof: Should be clear (see convention 2.5(1)).

Remark: Concerning 2.9(2) we can say something also in the case $b_0 \neq a_0$ but there was no real need.

2.10 Definition: For a σ-system $\mathfrak{B}$, and $r, n, m \in \mathbb{N}$ we define $\mathfrak{B}^n_{r,m}$ as the expansion of $\mathfrak{B}$ by the relations $R^\ell_{\mathbf{t}} = \{\bar{a} : \bar{a}$ is a $(\mathfrak{B}, r')$-component , $\mathbf{t} = \text{bth}^n_{r'}(\bar{a}, \mathfrak{B})$, $[\ell = 1 \Rightarrow \bar{a} \subseteq M]$ and $[\ell = 2 \Rightarrow \bar{a} \subseteq I]\}$ for each $\mathbf{t} \in \text{BTH}^n_{r'}(m', \sigma)$ $m' \leq m + n$, $r' \leq f^{(3)}_n(r)$ and $\ell \in \{1, 2\}$. We let $I^n_{r,m}[\mathfrak{B}] = \mathfrak{B}^n_{r,m} \restriction I$. Writing $\bar{r} = \langle r_\ell : \ell < k \rangle$ we mean $\max(\bar{r})$. Writing $\mathfrak{B}^n_{\bar{r},\bar{m}}$ means the common expansion of $\mathfrak{B}^n_{r_\ell, m_\ell}$ for $\ell < \ell g(\bar{r}) = \ell g(\bar{m})$ if $\ell g(\bar{r}) = 0$ we mean $\mathfrak{B}^n_{0,0}$ (we could alternatively use $\mathfrak{B}^n{}_{\max(\bar{r}),\, \max(\bar{m})}$, make little difference). Writing $\leq r$ we mean for every $r' \leq r$.

2.11 Claim: $I^{n+1}_{\bar{r},\bar{m}}[\mathfrak{B}]$ essentially expands $I^n_{\bar{r}',\bar{m}'}[\mathfrak{B}]$ when

(ii) $\bar{r}' = r \char`^ \langle 0 \rangle$, $\bar{m}' = \bar{m} \char`^ \langle 1 \rangle$ or

(ii) $\bar{r}' \leq f^{(2)}_n(\bar{r})$, $\bar{m}' \leq \langle m_i + 1 : i < \lg(\bar{m}) \rangle$

(essentially expand means that every predicate in the latter is equivalent to a quantifier free formula in the former, the function giving this is the scheme of expansion).

Proof: Should be clear by (i) of (β) of 2.8(1).

2.12 Definition: For a system $\mathfrak{B}$ and $n, m \in \mathbb{N}$ and $\bar{r} = \langle r_\ell : \ell < m \rangle$ such that $r_\ell \in \mathbb{N}$ and $\bar{a} = \langle a_0, \ldots, a_{k-1} \rangle \subseteq I^{\mathfrak{B}}$

1) We say $\bar{a}$ is $(n, \bar{r})$-sparse for $\mathfrak{B}$ if $\bar{r} = \langle r_\ell : \ell < m \rangle$ and

$$\ell < k < m \text{ implies } N^+_{f_n(r_\ell)}(a_\ell) \cap N^+_{f_n(r_k)}(a_k) = \emptyset$$

moreover (slightly stronger)

$$d(a_\ell, a_k) \geq f_n(r_\ell) + f_n(r_k) + 1.$$

2) We define $\mathrm{uth}^n_{\bar r}(\bar a, \mathfrak{B})$ for an $(n,\bar r)$-sparse $\bar a \subseteq \mathfrak{B}$, by induction on n:
$\mathrm{uth}^0_{\bar r}(\bar a, \mathfrak{B}) = \mathrm{th}^0(\bar a, \mathfrak{B})$ and $\mathrm{uth}^{n+1}_{\bar r}(\bar a, \mathfrak{B}) = \langle \mathbf{t}_0, \mathbf{t}_1 \rangle$ where:

$$\mathbf{t}_0 = \{\langle \bar s, \mathrm{uth}^n_{\bar s}(\bar a, \mathfrak{B})\rangle : \bar s \le f^{(2)}_n(\bar r)\}$$

(see 2.5, $f^{(2)}_n(\bar r) = \langle f_n(f_n(r_\ell)) : \ell < \ell g(\bar a)\rangle$, remember $f^{(2)}_n(\bar r) \le f_{n+1}(\bar r)$)

$$\begin{aligned}\mathbf{t}_1 = \{\langle \bar s, \mathrm{uth}^n_{\bar s ^\frown \langle 0\rangle}(\bar a ^\frown \langle c\rangle, \mathfrak{B})\rangle :\ & \bar s \le \bar r \text{ (i.e. } \bigwedge_\ell s_\ell \le r_\ell) \text{ and } c \in I^{\mathfrak{B}} \text{ and} \\ & \bar a ^\frown \langle c \rangle \text{ is } (n, \bar s ^\frown \langle 0\rangle)\text{-sparse i.e.} \\ & N^+_{f_n(0)}(c) \text{ is disjoint from } N^+_{f_n(s_\ell)}(a_\ell) \text{ for } \ell < \ell g(\bar a)\}.\end{aligned}$$

3) $\mathrm{UTH}^n_{\bar r}(m,\sigma)$ is the set of formally possible $\mathrm{uth}^n_{\bar r}(\bar a, \mathfrak{B})$ ($\bar a$ is $(n,\bar r)$-sparse for $\mathfrak{B}$ of length m, $\mathfrak{B}$ a σ-system, etc.).

2.13 Claim: 1) For every $\mathbf{t} \in \mathrm{UTH}^n_{\bar r}(m,\sigma)$ there is a formula $\varphi = \varphi(x_0,\dots,x_{m-1}) \in \mathcal{L}_{\sigma_{f^*_n(\bar r)}}$ where $f^*_n(\bar r) = \min\{m : \text{if } \ell_1 < \ell_2 < \ell g(\bar a) \text{ then } m > f_n(r_{\ell_1}) + f_n(r_{\ell_2})\}$ of quantifier depth n such that: For a σ-system $\mathfrak{B}$ and $(n,\bar r)$-sparse $\bar a = \langle a_0,\dots,a_{m-1}\rangle \subseteq I^{\mathfrak{B}}$, we have: $\mathfrak{B} \models \varphi[a_0,\dots,a_{m-1}]$ <u>iff</u> $\mathbf{t} = \mathrm{uth}^n_{\bar r}(\bar a, \mathfrak{B})$ (and being $(n,\bar r)$-sparse is equivalent to some quantifier free formula).
2) In Definition 2.12 only $\mathfrak{B} \restriction I$ matters and in part (1), the quantifications are on I only.
3) If $\mathfrak{B}'$ essentially expands $\mathfrak{B}$ then from (the scheme of expansion and) $\mathrm{uth}^n_{\bar r}(\bar a, \mathfrak{B}')$ we can compute $\mathrm{uth}^n_{\bar r}(\bar a, \mathfrak{B})$.

Proof: Should be clear.

2.14 Main Lemma: Let σ be a system-vocabulary; if $\otimes_0$ holds then $\otimes_n$ holds for every n where:
$\otimes_n$ there are functions $F_{n,\bar r,\bar m}$, for $\bar r = \langle r_\ell : \ell < k\rangle$, $\bar m = \langle m_\ell : \ell < k\rangle$, where $k, r_\ell, m_\ell \in \mathbb{N}$, such that:
$(*)$ <u>if</u> $\mathfrak{B}$ is a σ-system, $\bar a = \bar a^0 {}^\frown \dots {}^\frown \bar a^{k-1}$, $\bar a^\ell$ an $(\mathfrak{B}, r_\ell)$-component, $\langle a^\ell_0 : \ell < k\rangle$ is $(n,\bar r)$-sparse and $\bar m = \langle \ell g(\bar a^\ell) : \ell < k\rangle$ <u>then</u> letting $\mathbf{t}_\ell \overset{\text{def}}{=} \mathrm{bth}^n_{r_\ell}(\bar a^\ell, \mathfrak{B})$ (for $\ell < k$) and $\mathbf{t} \overset{\text{def}}{=} \mathrm{uth}^n_{\bar r}\Big(\langle h(a^\ell_0) : \ell < k\rangle, I^n_{\bar r,\bar m}[\mathfrak{B}]\Big)$ we have $\mathrm{th}^n_0(\bar a, \mathfrak{B}) = F_{n,\bar r,\bar m}(\mathbf{t}, \mathbf{t}_0, \dots, \mathbf{t}_{k-1})$.
$(**)$ $F_{n,\bar r,\bar m}$ is recursive in its variables, $n, \bar r, \bar m$ and the functions $F_{0,\bar r',\bar m'}$ where $\bar m' = \langle m'_i : i < k'\rangle$, $k' \ge k$, $\bar r' \le f^*_n(\bar r)$ (see 2.13(1)) and for $i < k$ we have $m'_i \ge m_i$ and $\sum_{i<k}(m'_i - m_i) + \sum_{i=k}^{k'-1} m'_i \le n$.

2.14A Remark: Why we use th^n_0 and not th^n_r (in the conclusion of $(*)$)? We can if we assume "$d(x,y) \le s \,\&\, x \in I \,\& y \in I$" is an atomic formula for $\mathfrak{B}$ for $s \le r$.

Proof: We prove this by induction on n, (for all $\bar r, \bar m$), so for $n = 0$ clearly $\otimes_n$ holds. So assume $\otimes_n$ and we shall prove $\otimes_{n+1}$. We shall now describe a value $\mathbf{t}$ computed from $\text{bth}^{n+1}_{r_\ell}(\bar a^\ell, \mathfrak{B})$ (for $\ell < k$) and $\text{uth}^{n+1}(\langle h(a^\ell_0) : \ell < k\rangle, I^{n+1}_{\bar r,\bar m}[\mathfrak{B}])$. Our intention is that $\mathbf{t} = \text{th}^{n+1}(\bar a, \mathfrak{B})$. Remember $\mathbf{t} = \{\text{th}^n(\bar a{}^\frown\langle c\rangle, \mathfrak{B}) : c \in \mathfrak{B}\}$.

Now $\mathbf{t}$ will be the union of two sets.

We use an informal description as it is clearer.

First Part: The set of $\text{th}^n(\bar a{}^\frown\langle c\rangle, \mathfrak{B})$, where for some $\ell(*) < k, c \in N^+_{f^{(2)}_n(r_{\ell(*)})}(a^{\ell(*)}_0)$.
Why can we compute this (using the induction hypothesis of course), for each such c we have: let $\bar b^\ell = \bar a^\ell$ if $\ell < k, \ell \ne \ell(*)$ and let $\bar b^{\ell(*)} = \bar a^{(\ell(*))\frown}\langle c\rangle$ (i.e. $b^{\ell(*)}_{m_{\ell(*)}} = c$); next r'_ℓ is: r_ℓ for $\ell \ne \ell(*)$ and $f^{(2)}_n(r_{\ell(*)})$ if $\ell = \ell(*)$. Necessarily for $\ell < k$, m^i_ℓ is m_ℓ if $\ell \ne \ell(*)$ and $m_{\ell(*)} + 1$ if $\ell = \ell(*)$. Now:

(α) $N^+_{f_n(r'_\ell)}(b^\ell_0)$ for $\ell < k$ are pairwise disjoint, moreover for $\ell(1) < \ell(2) < k$, $d(b^{\ell(1)}_0, b^{\ell(2)}_0) > f_n(r'_{\ell(1)}) + f_n(r'_{\ell(2)})$.
[Why? For $\ell = \ell(*)$ remember that for every r: $f_n(f^{(2)}_n(r)) = f^{(3)}_n(r)) \le f_{n+1}(r)$ and for $\ell \ne \ell(*)$ remember that for every r: $f_n(r) \le f_{n+1}(r)$ so in all cases clearly $f_n(r'_\ell) \le f_{n+1}(r_\ell)$ and $b^\ell_0 = a^\ell_0$. So if $\ell(1) \ne \ell(2)$ are $< k$, $N^+_{f_n(r'_{\ell(1)})}(a^{\ell(1)}_0) \cap N^+_{f_n(r'_{\ell(2)})}(b^{\ell(2)}_0) = \emptyset$ as: if $\ell(1) \ne \ell(*)$, $\ell(2) \ne \ell(*)$, trivial. Otherwise without loss of generality $\ell(2) = \ell(*)$, as $\bar a$ satisfies the assumption of $(*)$ ($(*)$ is from the lemma) $\bar a$ is $(n+1, \bar r)$-sparse hence (see Def. 2.12(1)) $d(a^{\ell(1)}_0, a^{\ell(*)}_0) > f_{n+1}(r_{\ell(1)}) + f_{n+1}(r_{\ell(2)})$, hence $d(b^{\ell(1)}_0, b^{\ell(*)}_0) = d(a^{\ell(1)}_0, a^{\ell(*)}_0) > f_{n+1}(r_{\ell(1)}) + f_{n+1}(r_{\ell(2)}) \ge f_n(r_{\ell(1)}) + f_n(f^{(2)}_n(r_{\ell(*)})) = f_n(r'_{\ell(1)}) + f_n(r'_{\ell(*)})$), as required.]

(β) $\bar b^\ell \subseteq N^+_{r'_\ell}(b^\ell_0)$, so $\bar b^\ell$ is an r'_ℓ-component.
[Why? when $\ell = \ell(*)$ as $r_\ell \le f^{(2)}_n(r_\ell) = r'_\ell$ and assumption on c; for $\ell \ne \ell(*)$ trivial.]

(γ) We can compute $\text{bth}^n_{r'_\ell}(\bar b^\ell, \mathfrak{B})$ for $\ell \ne \ell(*)$
[Why? by monotonicity properties of bth i.e. by 2.9(2) (that is clause (i) i.e. $\mathbf{t}_0$ from Def. 2.8(1)(β))];

(δ) We can compute the set of possibilities of $\text{bth}^n_{r'_{\ell(*)}}(\bar b^{\ell(*)}, \mathfrak{B})$.
[Why? those possibilities are listed in bth^{n+1} (see $\mathbf{t}_2$ in Definition 2.8(1)(β) mainly clause(ii)).]

(ε) We can compute $\mathrm{uth}^n_{\bar r'}(\langle h(b^\ell_0) : \ell < k\rangle, I^n_{r,\bar m'}[\mathfrak{B}])$.
[As we can compute $\mathrm{uth}^n_{\bar s}(\langle h(b^\ell_0) : \ell < k\rangle, I^{n+1}_{\bar r,\bar m}[\mathfrak{B}])$ for $\bar s \le f^{(2)}_n(\bar r)$ by the definition of $\mathbf{t}_0$ in Def. 2.12(2), choose $\bar s = \bar r'$; now, by 2.11, $I^{n+1}_{\bar r,\bar m}[\mathfrak{B}]$ essentially expand $I^n_{\bar r',\bar m'}[\mathfrak{B}]$ (see clause (ii) there) hence by 2.13(3) we can get the required object.]

(ζ) We can compute the set of possibilities of $\mathrm{th}^n(\bar b^0{}^\frown \dots {}^\frown \bar b^{k-1}, \mathfrak{B})$ gotten as above for fixed $\ell(*) < k$, all $c \in N^+_{f^{(2)}_n(r_{\ell(*)})}(a^{\ell(*)}_0)$.
[Why? by (α) - (ε) above and $\otimes_n$.]

(ξ) We can compute $\{\mathrm{th}^n(\bar a^\frown c, \mathfrak{B}) : \text{for some } \ell(*) < k,\ c \in N^+_{f^{(2)}_n(r_{\ell(*)})}(a^{\ell(*)}_0)\}$

Second Part: The set of $\mathrm{th}^n(\bar a^\frown\langle c\rangle, \mathfrak{B})$ where for each $\ell < k$, $c \notin N^+_{f^{(2)}_n(r_\ell)+f_n(0)}(a^\ell_0)$.

Why can we compute this? for such a c we can let $k' = k+1$, $\bar a^k = \langle c\rangle$ (so $c = a^k_0$, $m_k = 1$). Let $r'_\ell = r_\ell$, (for $\ell < k$) and $r'_k = 0$. Let $\bar m'$ be $\bar m^\frown\langle 1\rangle$. Now

(α) $N^+_{f_n(r'_\ell)}(a^\ell_0)$ for $\ell \le k$ are pairwise disjoint and $d(a^{\ell(1)}_0, a^{\ell(2)}_0) > f_n(r'_{\ell(1)}) + f_n(r'_{\ell(2)})$.
[Why? as $f_n(r) \le f_{n+1}(r)$ for $\ell(1) < \ell(2) < k$ this is trivial, and for $\ell(1) = \ell < k = \ell(2)$ we have $d(a^{\ell(1)}_0, a^{\ell(2)}_0) = d(a^\ell_0, a^k_0) = d(a^\ell_0, c) > f^{(2)}_n(r_\ell) \ge f_n(r_\ell) + f_n(0) = f_n(r'_\ell) + f_n(r'_k)$, as required].

(β) $\bar a^\ell \subseteq N^+_{r'_\ell}(a^\ell_0)$, so $\bar a^\ell$ is an r'_ℓ-component.
[Why? for $\ell < k$ as $r'_\ell = r_\ell$, for $\ell = k$ as $\bar a^k = \langle a^k_0\rangle (= \langle c\rangle)$].

(γ) we can compute $\mathrm{bth}^n_{r'_\ell}(\bar a^\ell, \mathfrak{B})$ for $\ell < k$
[Why? by monotonicity properties of bth i.e. 2.9(2)].

(δ) we can compute the possibilities for pairs $(\mathbf{t}', \mathbf{t}'')$ where $\mathbf{t}' = \mathrm{bth}^{n+1}_0(\langle c\rangle, \mathfrak{B})$ and $\mathbf{t}'' = \mathrm{uth}^n_{\bar r'}(\langle h(a^\ell_0) : \ell < k\rangle^\frown\langle h(c)\rangle, I^{n+1}_{\bar r,\bar m}[\mathfrak{B}])$.
[Why? straightforward; by the definition of uth^{n+1}, i.e. $\mathbf{t}_1$ of Def. 2.12(2) and $I^n_{\bar r,\bar m'}[\mathfrak{B}]$].

(ε) in (δ) we can replace $I^{n+1}_{\bar r,\bar m}[\mathfrak{B}]$ by $I^n_{\bar r',\bar m'}[\mathfrak{B}]$.
[Why? by 2.11 and 2.13(3).]

By the induction hypothesis this is enough.

Why the union of the two parts is $\mathrm{th}^n(\bar a, \mathfrak{B})$*?*

Both obviously give subsets, and if c fails the first part then $\ell < k \Rightarrow c \notin N^+_{f^{(2)}_n(r_\ell)}(a^\ell_0)$. So $N^+_{f_n(0)}(c)$ is disjoint to such $N^+_{f_n(r_\ell)}(a^\ell_0)$ and moreover $d(a^\ell_0, c) > f_n(r_\ell) + f_n(0)$.

$\square_{2.14}$

2.15 Conclusion: For any system vocabulary σ, and (first order) sentence φ of quantifier depth n, given $F_{0,\bar{r},\bar{m}}$'s satisfying $\otimes_0$ of 2.14 we can compute numbers r, m and a sentence ψ_φ of quantifier depth n, (whose vocabulary is that of $I^n_{r,m}[\mathfrak{B}]$) such that:
$(*)$ if $\mathfrak{B}$ is a σ-system which satisfied $\otimes_0$ as exemplified by $\langle F_{0,\bar{r},\bar{m}} : \bar{r},\bar{m}\rangle$ then $\mathfrak{B} \models \varphi \Leftrightarrow I^n_{r,m}[\mathfrak{B}] \models \psi_\varphi$.

Proof: By 2.14 and 2.13(1) and 2.7(1) (and see 2.14A(2)).

2.16 Conclusion: Let τ be a vocabulary (finite for simplicity) including a binary relation $<$ and a unary relation P, and $\varphi \in \mathcal{L}_\tau$. Then we can compute an $m < \omega$, formulas $\varphi_i(x) \in \mathcal{L}_\tau$ for $i < m$ with $d(\varphi_i) \le d(\varphi)$ and a sentence $\psi_\varphi \in \mathcal{L}_{\tau^*_1}$ with $d(\psi_\varphi) \le d(\varphi)$ where $\tau^*_1 \overset{\text{def}}{=} \{<\} \cup \{P_i : i < m\}$ ($m \in \mathbb{N}$ computable from φ, each P_i a unary predicate) satisfying the following:

$(*)$ Assume M is a finite τ-model, $<^M$ a linear order, $P \in \tau$ is unary, such that: if $R \in \tau \setminus \{<\}$ and $\bar{a} = \langle a_0, \dots, a_{n(R)-1}\rangle \in R^M$, then $P^M \cap [\min \bar{a}, \max \bar{a}]_M$ has at most one member.
Define $I[M] = (<^M, P^M, \dots, P_\ell \dots)_{\ell<m}$ where $P_\ell = \Big\{a \in P^M : \varphi_\ell[a]$ is satisfied in $M{\restriction}\{x : x \le a, |[x,a]_M \cap P^M| \le 3^{d[\varphi]}$ or $a \le x$, $|[a,x]_M \cap P^M| \le 3^{d[\varphi]}\}\Big\}$
($d[\varphi]$ in the quantifier depth of φ).
Then $M \models \varphi \Leftrightarrow I[M] \models \psi_\varphi$.

Proof: Should be clear from 2.15.

Remark: Concerning 2.16 we can deduce it also from §4.

§3 Proof of the Main Theorem

3.1 Definition: 1) For a finite set $J \subseteq I$ let $\mathrm{spr}(I,J)$ be the set of pairs $(Q^{\mathrm{no}}, Q^{\mathrm{yes}})$, where $Q^{\mathrm{no}} \subseteq I$, $Q^{\mathrm{yes}} \subseteq I$, $Q^{\mathrm{no}} \setminus J = Q^{\mathrm{yes}} \setminus J$, $|Q^{\mathrm{yes}} \vartriangle Q^{\mathrm{no}}| = 1$ ($A \vartriangle B$ is the symmetric difference). Let $\mathrm{spr}(I) = \mathrm{spr}(I,I)$.
2) For finite $J \subseteq I$ let $\mu^*(I,J)$ be the following distribution on $\mathrm{spr}(I,J)$; it is enough to describe a drawing:
first choose $Q^{\mathrm{no}} \subseteq I$ (all possibilities with probability $1/2^{|I|}$)
then choose $s \in J$ (all possibilities with probability $1/|J|$)
finally let

$$Q^{\mathrm{yes}} = \begin{cases} Q^{\mathrm{no}} \cup \{s\} & \textit{if } s \notin Q^{\mathrm{no}} \\ Q^{\mathrm{no}} \setminus \{s\} & \textit{if } s \in Q^{\mathrm{no}} \end{cases}$$

We write $\mu^*(I)$ for $\mu^*(I,I)$.

3.1A Remark: Note that the distribution $\mu^*(I,J)$ is symmetric for $Q^{\rm yes}, Q^{\rm no}$.

3.2 Definition: 1) For a linear order I and $J \subseteq I$ let $\mathrm{npr}(I,J) = \{(Q^{\rm no}, Q^{\rm yes}) : Q^{\rm no} \subseteq Q^{\rm yes} \subseteq I$ and $Q^{\rm no} \setminus J = Q^{\rm yes} \setminus J$ and $|Q^{\rm yes} \setminus Q^{\rm no}| = 1\}$.
2) If I is a set of natural numbers we use the usual order.
3) If $J = I$ we write $\mathrm{npr}(I)$.
4) Any (I,J) is isomorphic to some (n, J') so we can use such pairs above.
5) Let $\mu^{**}(I,J)$-be the distributions $\mu^*(I,J)$ restricted to the case $Q^{\rm no} \subseteq Q^{\rm yes}$ (i.e. to $\mathrm{npr}(I,J)$).

3.3 Claim: Let $m, d \in \mathbb{N}$ be given, $\tau = \{<\} \cup \{P_i : i < m\}$, P_i a unary predicate, $K = \{M : M$ a τ-model, $<^M$ a linear order$\}$.
1) For every $\varepsilon \in \mathbb{R}^+$ for every k large enough, for the distribution $\mu^*(k)$ on $\mathrm{spr}(k) = \{(Q^{\rm no}, Q^{\rm yes}) : Q^{\rm no}, Q^{\rm yes} \subseteq k, |Q^{\rm yes} \vartriangle Q^{\rm no}| = 1\}$ we have: <u>if</u> $M_i^{\mathfrak{t}} \in K$, for $i < k$, $\mathfrak{t} \in \{\text{yes no}\}$, and we choose $\mu^*(k)$-randomly $(Q^{\rm no}, Q^{\rm yes}) \in \mathrm{spr}(k)$, <u>then</u>
$(*)$ the probability of $\mathrm{th}^d(\sum_{i<k} M_i^{\text{if } (i \in Q^{\rm yes})}) = \mathrm{th}^d(\sum_{i<k} M_i^{\text{if } (i \in Q^{\rm no})})$ is at least $1 - \varepsilon$ (th^d is defined as in 2.6(1) considering a model as a system by 2.2(4)).
2) Also if we first choose $(Q_u^{\rm no}, Q_u^{\rm yes}) \in \mathrm{spr}\Big([[\frac{k+1}{2}], k)\Big)$ as above and then (possibly depending on the result) make a decision on a choice of $Q_d^{\rm no} = Q_d^{\rm yes} \subseteq [0, [\frac{k+1}{2}])$ and let $Q^{\rm yes} = Q_d^{\rm yes} \cup Q_u^{\rm yes}$, $Q^{\rm no} = Q_d^{\rm no} \cup Q_u^{\rm no}$ <u>then</u> $(*)$ above still holds.
3) If we choose $Q^{\rm yes} \subseteq k$ such that $|Q^{\rm yes}| = [\frac{k+1}{2}]$ and then $Q^{\rm no} \subseteq Q^{\rm yes}$, $|Q^{\rm no}| + 1 = |Q^{\rm yes}|$ (all possibilities with the same probabilities) <u>then</u> with probability tending to 1 with k going to ∞ we get that $(*)$ (of 3.3(1) above) holds.
4) The parallel of (1) holds for $\mu^{**}(k)$, $\mathrm{npr}(k)$.

Before proving 3.3 we define and note:

3.4Definition: For the τ, K, d from 3.3, let ζ_k be the maximal real in $[0,1]$ such that: if $M_i^{\mathfrak{t}} \in K$ for $i < r$, $\mathfrak{t} \in \{\text{yes}, \text{no}\}$ (where $r \in \mathbb{N}$), $J \supseteq \{i < r : \mathrm{th}^d(M_i^{\rm no}) \neq \mathrm{th}^d(M_i^{\rm yes})\}$ has k elements (and $J \subseteq I = r = \{0, \ldots, r-1\}$), <u>then</u> $\frac{k}{r}(1 - \zeta_k) \geq \mathrm{Prob}_{\mu^*(I,J)}\Big(\mathrm{th}^d(\sum_{i<r} M_i^{\text{if } (i \in Q^{\rm no})}) \neq \mathrm{th}^d(\sum_{i<r} M_i^{\text{if } (i \in Q^{\rm yes})})\Big| (Q^{\rm no}, Q^{\rm yes}) \in \mathrm{spr}(I,J)\Big)$. Let $\xi_k = 1 - \zeta_k$.

3.4A Observation: 1) ζ_k is well defined; and $\zeta_k \leq \zeta_{k+1}$
2) An alternative definition of ζ_k is that it is the maximum real in $[0,1]$ satisfying: $(1 - \zeta_k)$ is not smaller than the relative-probability of $\mathrm{th}^d\Big(\sum_{i<r} M_i^{\text{if}(i \in Q^{\rm no})}\Big) \neq$

$\text{th}^d\Big(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})}\Big)$ for the probability distribution $\mu^*(I)$, under the assumption $Q^{\text{no}} \vartriangle Q^{\text{yes}} \subseteq J$.
3) Without loss of generality in 3.4, $r \le 2k+1$ (and even $r = k$).

Proof: Note that the number of possible $\langle \text{th}^d(M_i^{\mathbf{t}}) : i, \mathbf{t}\rangle$ is finite.
1) By (2). First draw $i \in J$ (equal probability) and then use $\text{spr}(I, J \setminus \{i\})$. Alternatively let $J = \{i_0, \dots, i_{\ell(*)-1}\}$, where $0 \le i_0 < \dots < i_{\ell(*)-1} < r$, $I = \{0, \dots, r-1\}$ so $|J| = \ell(*)$. Let $J_{\ell(*),\ell} \overset{\text{def}}{=} J \setminus \{i_\ell\}$ for $\ell < \ell(*)$. Clearly
$(*)_1$

$$\text{Prob}_{\mu(I,J)}\Big(\text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}) \ne \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\Big|(Q^{\text{no}}, Q^{\text{yes}}) \in \text{spr}(I,J)\Big)$$

$$= \frac{1}{\ell(*)} \sum_{\ell<\ell(*)} 2^{-(r-1)} |\{Q \subseteq I\backslash\{i_\ell\} : \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q)}) \ne \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q\cup\{i_\ell\})})\}|.$$

Hence if $\ell(*) = k+1$ then
$(*)_2$

$$\text{Prob}_{\mu(I,J)}\Big(\text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}) \ne \text{th}^d(\sum_{i<r} (M_i^{\text{if}(i\in Q^{\text{yes}})})\Big|(Q^{\text{no}}, Q^{\text{yes}}) \in \text{spr}(I,J)\Big)$$

$$= \frac{1}{k+1} \sum_{\ell\le k} \text{Prob}_{\mu(I,J_{k+1,\ell})}\Big(\mathcal{E}_\ell\Big|(Q^{\text{no}}, Q^{\text{yes}}) \in \text{spr}(I, J_\ell)\Big),$$

where

$$\mathcal{E}_\ell = \{(Q^{\text{no}}, Q^{\text{yes}}) \in \text{spr}(I,J) : \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}) = \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\}.$$

Now compute.
2) Should be clear.
3) By addition theory i.e. 2.7(7) and by 2). $\square_{3.4A}$

3.5 Definition: Let $c \in \mathbb{N}$ be the number of members in $\text{TH}^d(0,\tau)$ (the set of formally possible $th^d(\langle\rangle, M)$, $M \in K$, see 2.6(3)). Let $k_0 \in \mathbb{N}$ be such that $k_0 \to (3^{d+8})^2_{c^2}$ (exists by Ramsey theorem).

3.6 Observation:
$(*)_1 \quad \zeta_{k_0} \ge \frac{1}{(k_0 2^{k_0})}$.
[Why? Let r, $M_i^{\mathbf{t}}$, and J, $|J| = k_0$ be given. First draw $Q^{\text{yes,no}} \cap (r \setminus J)$ and

assume they are equal.
Now (by 3.4A(2)) it is enough to prove that now the probability of the equality i.e. of

$$\text{th}^d\Big(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}\Big) = \text{th}^d\Big(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\Big)$$

is $\geq 1/(k_0 2^{k_0})$, assuming $Q^{\text{no}} \vartriangle Q^{\text{yes}} = \{j\} \subseteq J$. For $i < j$ from J let $\langle \mathbf{t}^{\text{no}}_{i,j}, \mathbf{t}^{\text{yes}}_{i,j}\rangle$ be $\langle \text{th}^d(M_i^{\text{no}} + M_{i+1}^{\text{no}} + \ldots + M_{j-1}^{\text{no}}), \text{th}^d(M_i^{\text{yes}} + M_{i+1}^{\text{no}} + M_{i+2}^{\text{no}} + \ldots + M_{j-1}^{\text{no}})\rangle$. So by the choice of k_0 we can find $J' \subseteq J, |J'| = 3^{d+8}$ and $\langle \mathbf{t}_0, \mathbf{t}_1\rangle$ such that for $i < j$ in J', we have $\langle \mathbf{t}^{\text{no}}_{i,j}, \mathbf{t}^{\text{yes}}_{i,j}\rangle = \langle \mathbf{t}_0, \mathbf{t}_1\rangle$. Let $J' = \{i_\ell : \ell < 3^{d+8}\}$. For each $j \leq 3^{d+8}$ let $Q_i = \{i_m : m < j\}$. By addition theory for (first order theory) linear order, (that is 2.7(6)) for $\ell \stackrel{\text{def}}{=} [3^{d+8}/2]$, $\text{th}^d(\sum_i M_i^{\text{if }(i\in Q_\ell)}) = \text{th}^d(\sum_i M_i^{\text{if }(i\in Q_{\ell+1})})$. So the probability for equality is at least the probability of $Q^{\text{no}} = Q_\ell$, $Q^{\text{yes}} = Q_{\ell+1}$ which is $\geq 1/(k_0 2^{k_0})$]. $\square_{3.6}$

3.7 Observation: For every ℓ, $k > 0$ we have

$$\xi_{k\ell} \leq \xi_k\Big(\sum_{j\leq \ell-1} [\binom{\ell-1}{j} \xi_k^j (1-\xi_k)^{\ell-1-j}]\xi_j\Big).$$

Proof: Let us be given r, $M_i^{\mathbf{t}}$ for $i < r$ and J as in $(*)$ so $|J| = k\ell$. Choose $\langle I_j : j < \ell\rangle$ a partition of $I \stackrel{\text{def}}{=} r = \{0,\ldots,r-1\}$ to intervals such that for $j < \ell$, $J_j \stackrel{\text{def}}{=} J \cap I_j$ has exactly k members. Now first draw $Q_j^{\text{no}} \subseteq I_j$ for $j < \ell$ (with equal probabilities), second draw $s_j \in J_j$ for $j < \ell$ (with equal probabilities) and third draw $R^{\text{no}} \subseteq \{0,\ldots,\ell-1\}$ (with equal probabilities) and fourth draw $j(*) < \ell$ (with equal probabilities). Let $Q_j^{\text{yes}} = Q_j^{\text{no}} \vartriangle \{s_j\}$, and $R^{\text{yes}} = R^{\text{no}} \vartriangle \{j(*)\}$, and $Q^{\text{no}} = \bigcup_{j<\ell} Q_j^{\text{if}(\ell\in R^{\text{no}})}$, $Q^{\text{yes}} = \bigcup_{j<\ell} Q_j^{\text{if}(\ell\in R^{\text{yes}})}$.

Easily $(Q_j^{\text{no}}, Q_j^{\text{yes}})$ was chosen by the distribution $\mu^*(I_j, J_j)$ and $(R^{\text{no}}, R^{\text{yes}})$ was chosen by the distribution $\mu^*(\{0,\ldots,\ell-1\})$ and $(Q^{\text{no}}, Q^{\text{yes}})$ was chosen by the distribution $\mu^*(I, J)$. Hence it is enough to prove:

$(*)$ $\text{Prob}\Big(\text{th}^d(\sum_{i<r} M^{\text{if}(i\in Q^{\text{no}})}) \neq \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\Big)$
$\leq \xi_k(\sum_{j\leq\ell-1}[\binom{\ell-1}{j}\xi_k^j(1-\xi_k)^{\ell-1-j}]\xi_j)$.

For $j < \ell$ let $N_i^{\mathbf{t}} = \sum_{i\in I_j} M^{\text{if}(i\in Q_j^{\mathbf{t}})}$ and let $p_i \in [0,1]_{\mathbb{R}}$ be $\text{Prob}\Big(\text{th}^d(N_j^{\text{no}}) \neq \text{th}^d(N_j^{\text{yes}})\Big)$ and let $A = \{j : \text{th}^d(N_j^{\text{no}}) \neq \text{th}^d(N_j^{\text{yes}})\}$, so the events "$j \in A$" are independent and $p_i = \text{Prob}(i \in A) \leq \xi_k$. Now if we make the first and second drawing only, we know A, $\langle N^{\mathbf{t}}_i : i < \ell, \mathbf{t} \in \{\text{no}, \text{yes}\}\rangle$ and modulo this, by the definition of $\xi_{|A|}$ we know

$$\text{Prob}\Big(\text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}) \neq \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\Big| \text{ after 1st and 2nd drawing}\Big) =$$

$$\text{Prob}\Big(\text{th}^d(\sum_{j<\ell} N_j^{\text{if}(j\in R^{\text{no}})}) \neq \text{th}^d(\sum_{j<\ell} N_j^{\text{if}(j\in R^{\text{yes}})})\Big| \text{ after 1st and 2nd drawing}\Big) \leq \xi_{|A|}.$$

As the events "$j \in A$" are independent we can conclude

$$\text{Prob}_{\mu^*(I,J)}\Big(\text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}) \neq \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\Big|(Q^{\text{no}}, Q^{\text{yes}}) \in \text{spr}(I,J)\Big)$$

$$\leq \sum_{j\leq \ell} \text{Prob}(|A| = j) \times \frac{j}{\ell} \times \xi_j = \sum_{j\leq\ell}\Big(\sum_{u\subseteq \ell, |u|=j} \prod_{m\in u} p_m \prod_{m<\ell, m\notin u}(1-p_m)\Big) \times \frac{j}{\ell}\xi_j.$$

Now looking at this as a function in $p_m \in [0, \xi_k]_{\mathbb{R}}$ for $m < \ell$, for some $\langle p^*_m : m < \ell\rangle$ we get maximal values, and as the function is linear, $p^*_m \in \{0, \xi_k\}$, and as $0 \leq \xi_j \leq \xi_{j+1} \leq 1$, necessarily $p^*_m = \xi_k$ so

$$\text{Prob}_{\mu^*(I,J)}\Big(\text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{no}})}) \neq \text{th}^d(\sum_{i<r} M_i^{\text{if}(i\in Q^{\text{yes}})})\Big|(Q^{\text{no}}, Q^{\text{yes}}) \in \text{spr}(I,J)\Big)$$

$$\leq \sum_{j\leq\ell}(\sum_{u\subseteq\ell, |u|=1}\prod_{m\in u} p^*_m \prod_{m<\ell, m\notin u}(1-p^*_m)) \times \frac{j}{\ell} \times \xi_j =$$

$$= \sum_{j\leq\ell}\binom{n}{j}(\xi_k)^j(1-\xi_k)^{\ell-j}(\frac{j}{\ell})\xi_j =$$

$$= \sum_{0<j\leq\ell}\binom{\ell}{j}(\xi_k)^j(1-\xi_k)^{\ell-j}\frac{j}{\ell}\xi_j =$$

$$= \sum_{j\leq\ell-1}[\binom{\ell}{j+1}\frac{(j+1)}{\ell}(\xi_k)^{j+1}(1-\xi_k)^{\ell-j-1}]\xi_j =$$

$$\xi_k \sum_{j\leq\ell-1}[\binom{\ell-1}{j}(\xi_k)^j(1-\xi_k)^{(\ell-1)-j}]\xi_j.$$

$\square_{3.7}$

3.8 Observation.
1) If ℓ, $k > 0$, $j_0 \leq \ell\xi_k$ then $\xi_{k\ell} \leq \xi_k(\frac{1+\xi_{j_0}}{2})$.
2) If $\xi_k \leq 1 - \frac{1}{m}$, $\ell > k/\xi_k$ then $\xi_{k\ell} \leq \xi_k(\frac{1+\xi_k}{2})$.

Proof: 1) As $\xi_j \geq \xi_{j+1}$ we have

$$\xi_{k\ell} \leq \xi_k \times (\sum_{j \leq \ell-1} [\binom{\ell-1}{j}(\xi_k)^j(1-\xi_k)^{\ell-1-j}]\xi_j)$$

$$\leq \xi_k \times \Big(\sum_{j<j_0}[\binom{\ell-1}{j}(\xi_k)^j(1-\xi_k)^{\ell-1-j}] \times 1+$$

$$\sum_{j \in [j_0,\ell-1)} [\frac{(\ell-1)!}{j!(\ell-j-1)!}(\xi_k)^j(1-\xi_k)^{\ell-1-j}] \times \xi_{j_0}\Big)$$

$$\leq \xi_k(\frac{1+\xi_{j_0}}{2}).$$

2) Follows. $\square_{3.8}$

3.8A Remark. Using "the binomial distribution approach normal distribution" and 3.6, clearly we get e.g.:

for every $\varepsilon > 0$, for some ℓ_ε, for every $\ell \geq \ell_\varepsilon$ we have

$$\xi_{k\ell} \leq \xi_k \xi_{(1-\varepsilon)\ell}.$$

3.9 Proof of 3.3: 1) By the definition of ζ_k and Observations 3.5, 3.6 we get that $\lim_{k\to\infty} \zeta_k = 1$ and we can finish easily.
2) Follows by (1) (and the addition theory see 2.7, particularly 2.7(7))
3) Similar proof and not used (e.g. imitate the proof of 3.6. First choose $j(*)$ then we have probability ζ_κ for equality there if the distribution is $\mu^*(I_{j(*)}, J_{j(*)})$, but the induced distribution is very similar to it).
4) Follows very easily. For spr(k) with probability $1/2$ we are choosing by npr(k).
$\square_{3.3}$

3.10 Proof of 1.4: Let a real $\varepsilon > 0$ and a sentence $\theta \in \mathcal{L}^{\text{fo}}_\tau$ (τ -from 1.3) be given. We shall define φ below (after $(*)_1$), and let $\psi = \psi_\varphi$, τ^*_1 (a vocabulary) and $m \in \mathbb{N}$ and $\varphi_i(x)$ for $i < m$ be defined as in 2.16 (for the φ here). Let d be the quantifier depth of ψ_φ (i.e. $d = d[\psi_\varphi]$). Let $k^* \in \mathbb{N}$ be large enough as in 3.3(4) (for the given ε, m and τ^*_1). Let k be $(2k^*+2)(3^{d[\theta]}+1)$ (we could have waived the $3^{d[\theta]}+1$). Now choose by induction on $r \leq k$, $m_r \in \mathbb{N}$ such that

$(*)$ (a) $0 = m_0$

(b) $m_r < m_{r+1} < \ldots$

(c) for any n,

$$\varepsilon/3 > \mathrm{Prob}_{\mu_n}\Big(M_n \models \bigvee_r (\exists x \le m_r + 1)(\exists y \ge m_{r+1} - 1)[xRy]\Big|M_n \in K_n\Big).$$

[Why this is possible? We choose m_r by induction on r. The probability in question is, for each fixed r, bounded from above by $\sum_{i<m_r}\sum_{j>m_{r+1}} p_{j-i}$, the sum is the tail of an (absolutely) convergent infinite sum so by increasing m_{r+1} we can make it $< \varepsilon/2^{r+2}$, this suffices].

Next we try to draw the model M_n in another way. Let $n > m_k$ be given; let $n^* = n+k^*+1$. Let $J = \{m_{(3^{d[\theta]+1})i} : 0 < i < 2k^*\}$, and $I = \{m_i : i < 2k^* 3^{d[\theta]+1}\}$. We first draw $\mathfrak{A}_n$, "a drunkard model $\mathfrak{A}_n$ for n". Drawing $\mathfrak{A}_n$ means:
laziness case=first case if $i < j < n^*$, $\bigvee_{r<k}[i \le m_r + 1 \& j \ge m_{r+1} - 1]$ then $\{i,j\}$ is non edge (no drawing).
normal case=second case: if $i < j < n^*$ are not in the first case but $\neg(\exists m \in I)[i \le m \le j]$ then we flip a coin and get $e_{i,j} \in \{\text{yes, no}\}$ with probability p_{j-i} (for yes).
drunkard case=third case: $i < j < n^*$ and no previous case apply; we make two draws. In one we get $e^1_{i,j} \in \{$ yes, no$\}$ with probability p_{j-i-1} (for yes) in the second we get $e^2_{i,j} \in \{$ yes, no$\}$ with probability p_{j-i} for yes (we may stipulate $p_0 = 0$).

Now for every $Q \subseteq J$ we define $M_Q[\mathfrak{A}_n]$: it is a model $(A^Q, <, P, R^Q)$ where A^Q is $\{0, \dots, n^* - 1\} \setminus Q$ (so $\|M_Q[\mathfrak{A}_n]\|$ is $n^* - |Q|$ and usually $|Q|$ will be k^* or $k^* + 1$)
$<$ is the usual order on I^Q
$P = \{m_r : r < k\}$
R^Q is† $\{(i,j) : \{i,j\} = \{i',j'\}, i' < j' < n^*$, and:

(a) (i',j') fall into the second case above and $e_{i',j'} =$ yes
or (b) (i',j') fall into the third case say $i \le m \le j$ and $m \in I$
(m is unique by " not first case") $m \in Q$, and $e^1_{i',j'} =$ yes
or (c) (i',j') fall into the third case say $i \le m \le j$ and $m \in I$
(m is unique by " not first case") $m \notin Q$ and $e^2_{i',j'} =$ yes$\}$.

We also define a model $N[\mathfrak{A}_n]$:
the universe: $\{0, \dots, n^* - 1\}$
relations: $<$ the usual order

$$R = \{\langle i,j\rangle : e_{i,j} = \text{yes},\ i < j\}$$

† We use i, j so that without loss of generality $i' < j'$.

$R^1 = \{\langle i,j\rangle : e^1_{i,j} = \text{yes},\ i<j\}$
$R^2 = \{\langle i,j\rangle : e^2_{i,j} = \text{yes},\ i<j\}$
$P = \{m_r : r \le k\}$ (on κ see above, before $(*)$)

Observe

$(*)_1$ in $(N[\mathfrak{A}_n],Q)$ we can define $M_Q[\mathfrak{A}_n]$ by q.f. formulas.

So for some first order φ depending on θ,τ but not on n (promised above in the beginning of the proof):

$(*)_2$ $M_Q[\mathfrak{A}_n] \models \theta$ iff $(N[\mathfrak{A}_n],Q) \models \varphi$.

By 2.16 (where $I[(N[\mathfrak{A}_n],Q)]$ is defined in 2.16 with M there standing for $(N[\mathfrak{A}_n],Q)$ here, so its set of elements is P) and the choice of ψ_φ we have:

$(*)_3$ $(N[\mathfrak{A}_n],Q) \models \varphi$ iff $I[(N[\mathfrak{A}_n],Q)] \models \psi_\varphi$.

Looking at the definition of $I[N[\mathfrak{A}_n],Q]$ in 2.16 without loss of generality

$(*)_4$ $I[(N[\mathfrak{A}_n],Q)] \models \psi_\varphi$ iff $(I[N[\mathfrak{A}_n]],Q) \models \psi_\varphi$.

Let $J = J^d \cup J^u$, J^d an initial segment, J^u an end segment, $|J^u| = k^*$, $|J^d| = k^*-1$. Now we define further drawing; let $\mu^{**}[J,J^u]$ be the distribution from 3.2 above on $\mathrm{npr}(J^u)$, and choose $(Q^u_0,Q^u_1) \in \mathrm{npr}(J^u)$ randomly by $\mu^{**}(J^u)$ then choose $Q^d_1 \subseteq J^d$ such that $|Q^d_1| = k^*+1-|Q^u_1|$ with equal probabilities, and let $Q^1 = Q^d_1 \cup Q^u_1$, and let $Q^d_0 = Q^d_1$, $Q^0 = Q^d_0 \cup Q^u_0$.

Now for $\ell \in \{0,1\}$ we make a further drawing: if $i<j$ is a pair from the first possibility (in the drawing of $\mathfrak{A}_n$), we flip a coin for ${}^*e^\ell_{i,j} \in \{\text{yes},\text{no}\}$ with probability $p_{|[i,j)\setminus Q^\ell|}$ for yes.

Let $M^\ell_{Q^\ell}$ be $(A^{Q^\ell},<,R^{Q^\ell} \cup \{(i,j),(j,i) : i<j \text{ and: } {}^*e^\ell_{i,j} = \text{yes}\})$ (it depends on the choice of $\mathfrak{A}_n$ and on the further drawing).

Now reflecting we see

$(*)_5$ for $\ell = 0,1$, the distribution of $M^\ell_{Q^\ell}$ is the same as that of $(K_{n+1-\ell},\mu_{n+1-\ell})$ (from Def. 1.3).

Hence

$(*)_6$ for $\ell=0,1$, $\mathrm{Prob}(M^\ell_{Q^\ell} \models \theta) = \mathrm{Prob}(M_{n+1-\ell} \models \theta \Big| M_{n+1-\ell} \in K_{n+1-\ell})$.

By the choice of m_r's

$(*)_7$ $\mathrm{Prob}(M^\ell_{Q^\ell} = M_{Q^\ell}[\mathfrak{A}_n])$ is $\ge 1-\varepsilon/3$.

By 3.3 above (used above: the drawing of (Q^u_0,Q^u_1) was randomly by $\mu^{**}(J^u)$).

$(*)_8$ the absolute value of the differences between the following is $\le \varepsilon/3$:

$\mathrm{Prob}([N[\mathfrak{A}_n]],Q^0) \models \psi_\varphi)$
$\mathrm{Prob}([N[\mathfrak{A}_n]],Q^1) \models \psi_\varphi)$.

So for $\ell = 0,1$:

(a) $\mathrm{Prob}(M_{n+\ell} \models \theta \Big| M_{n+\ell} \in K_{n+\ell}) = \mathrm{Prob}(M^\ell_{Q^\ell} \models \theta)$
[Why? by $(*)_6$.]

(b) $\mathrm{Prob}(M^\ell_{Q^\ell} = M_{Q^\ell}[\mathfrak{A}_n]) \geq 1 - \varepsilon/3$
[Why? by $(*)_7$.]

(c) $M_{Q^\ell}[\mathfrak{A}_n] \models \theta$ iff $(I[N[\mathfrak{A}_n]], Q^\ell) \models \psi_\varphi$
[Why? by $(*)_1 + (*)_2 + (*)_3 + (*)_4$.]

By (a)+(b)+(c) it suffices to prove that the probabilities of

$$\text{“}(I[N[\mathfrak{A}_n]], Q^\ell) \models \psi_\varphi\text{”}$$

for $\ell = 0, 1$ has difference $< \varepsilon/3$ but this holds by $(*)_8$. $\square_{1.4}$

§4 Generalized sums and Distortions:

We try here to explain the results on §2 as distorted generalized sums (and the connection with generalized sums) and later the connection to models with distance. First we present for background the definition and theorem of the generalized sum.

4.1 Definition: Let τ_0, τ_1, τ_2 be vocabularies of models. For a τ_0-model I (serving as an index model), τ_1-models, pairwise disjoint for simplicity $M_t (t \in I)$, and function F (explained below) we say that a τ_2-model M is the F-sum of $\langle M_t : t \in I \rangle$) in symbols $M = \oplus_F \{M_t : t \in I\}$ <u>if:</u>

(a) the universe $|M|$ of M is $\bigcup_{t \in I} |M_t|$ (if the M_t's are not pairwise disjoint: $\{(t,a) : t \in I, a \in M_t\}$) and we define $h_\alpha : M \to I$, $h(a) = t$ if $a \in M_t$ (if not disjoint $h(\langle t, a \rangle) = t$),

(b) if $t_1, \ldots, t_k \in I$ are pairwise distinct, $\bar a = \bar a_1{}^\frown \ldots {}^\frown \bar a_k$ and $\bar a_\ell \in M_{t_\ell}$ (finite sequences) *then*

$$\mathrm{tp}_{\mathrm{qf}}(\bar a, \emptyset, M) = F\Big(\mathrm{tp}_{\mathrm{qf}}(\langle t_1, \ldots, t_k \rangle, \emptyset, I), \mathrm{tp}_{\mathrm{qf}}(\bar a_1, \emptyset, M_{t_1}), \ldots, \mathrm{tp}_{\mathrm{qf}}(\bar a_k, \emptyset, M_{t_k})\Big).$$

Another way to say it is:

(b)$'$ if $a_0, \ldots, a_{k-1} \in M$ then $\mathrm{tp}_{\mathrm{qf}}(\langle a_0, \ldots, a_{k-1} \rangle, \emptyset, M) =$

$$F\Big(\mathrm{tp}_{\mathrm{qf}}(\langle h(a_\ell) : \ell < k \rangle), \ldots, \mathrm{tp}_{\mathrm{qf}}(\langle a_\ell : \ell < k, t = h(a_\ell) \rangle, \emptyset, M_t), \ \ldots\Big)_{t \in \{h(a_\ell) : \ell < k\}}.$$

4.2 Remark: 1) So the form of F is implicitly defined, it suffices to look at sequences $\bar{a}_0{}^\frown \ldots {}^\frown \bar{a}_k$ (in clause (b)) or $\langle a_0, \ldots, a_k \rangle$ of length the arity of τ_2 (if it is finite) i.e. maximum numbers of places for P predicates $P \in \tau_2$.
2) We can consider a generalization where the universe of M and equality are defined like any other relation.

4.3 The generalized Sum Theorem: In the notation above if $\mathbf{\Delta}_n$ is the set of formulas of quantifier depth n then we can compute from F the following function: like F in (b) (or (b)$'$) replacing $\mathrm{tp}_{\mathrm{qf}}$ by $\mathrm{tp}_{\mathbf{\Delta}_n}$.

4.4 Discussion: So looking at a sequence $\bar{a}$ from M, to find its [quantifier free] types we need to know two things:
(α) the [quantifier free] type of its restriction to each $h^{-1}(\{t\}) = \{b \in M : h(b) = t\}$ in M_t
(β) the [quantifier free] type of the sequence of parts, $\langle h(b_\ell) : \ell < k \rangle$ *in* I.

4.5 Definition: Now M is a d-distorted F-sum of $\{M_t : t \in I\}$ if

(a) d is a distance function on I.

(b) $|M|$ is the disjoint union of A_t $(t \in I)$ and for each t we have a model M_t with universe $\cup\{A_s : s \in I, d(s,t) \leq 1\}$ and

(c)$^+$ if $b_0, \ldots, b_{k-1} \in M$, then

$$\begin{aligned}&\mathrm{tp}_{\mathrm{qf}}(\langle b_0, \ldots, b_{k-1} \rangle, \emptyset, M) = \\ &F\Big(\mathrm{tp}_{\mathrm{qf}}(\langle h(b_0) \ldots, h(b_{k-1}), \emptyset, I), \ldots, \\ &\qquad \mathrm{tp}_{\mathrm{qf}}(\langle b_\ell : \ell < k, d(h(b_\ell), t) \leq 1), \emptyset, M_t), \ldots \Big)_{t \in \{h(b_m) : m < k\}}.\end{aligned}$$

4.6 Remark: Note: by $\langle b_\ell : \ell < k, \Pr(\ell) \rangle$ we mean the function g with domain $\{\ell < k : \Pr(\ell)\}$, satisfying $g(\ell) = b_\ell$).

4.7 Discussion: Our main Lemma 2.14, generalizes the generalized sum theorem, to distorted sum but naturally the distortion "expands" with the quantifier depth.

4.8 The Distorted Sum Generalized Lemma: In the notation above, for m let M^m_t be the model with
universe: $\cup\{A_s : d(s,t) \leq m\} \cup \{s \in I : d(s,t) \leq m\}$

relations: those of the M_s's i.e. for $R \in \tau_1$, a k-place predicate we let

$$Q_R^{t,m} = \{\langle s, a_1, \ldots, a_k\rangle : s \in I, d(s,t) \leq m+1, \langle a_1, \ldots, a_k\rangle \in R^{M_s}\}$$

(so $\{a_1, \ldots, a_k\} \subseteq \cup\{A_s : d(s,t) \leq 1\}$)

$$Q_d^{t,\ell,m} = \{(s_1, s_2) : d(s_1,t) \leq \ell \text{ and } d(s_2,t) \leq \ell,\ d(s_1,s_2) \leq m\},$$

$$Q_h^t = \{(s,a) : a \in A_s, s \in I, d(s,t) \leq m\}.$$

We define $I^{[m,n]}$ as the expansion of I by:
for $\ell \leq m$ and $\varphi \in \boldsymbol{\Delta}_n(\tau_0)$

$$Q_\varphi^\ell = \{t \in I : M_t^\ell \models \varphi\}.$$

Now there are functions F_n and a number $m(n) = 3^n$ computable from F (and n) such that:

$\otimes$ for $b_0, \ldots, b_{k-1} \in M$
$$\mathrm{tp}_{\boldsymbol{\Delta}_n}(\langle b_0, \ldots, b_{k-1}\rangle, \emptyset, M) = F_n\Big(\mathrm{tp}_{\boldsymbol{\Delta}_n}(\langle h(b_0), \ldots, h(b_{k-1})\rangle, \emptyset, I^{[m(n),n]}), \ldots,$$
$$\mathrm{tp}_{\boldsymbol{\Delta}_n}(\langle b_\ell : \ell < k, d(h(b_\ell),t) \leq m(n)\rangle, \emptyset, M_t^{m(n)}), \ldots\Big)_{t \in \{h(b_m) : m < k\}}.$$

4.9 Discussion: 1) Now if d is trivial:

$$d(x,y) = \begin{cases} 0 & x = y \\ \infty & x \neq y \end{cases}$$

then $M_t^n = M_t$, and 2.8 (the disorted generalized sum Lemma) becomes degenerated to 4.3 (the generalized sum Lemma), more exactly a variant.
2) Note that 4.8 improve on the result of §2 in $m(n)$ not depending on k. We can have this improvement in §2 and in §5.
3) To prove 4.8 given $b_0, \ldots, b_{k-1}$ and looking for y with

$$\mathbf{t} = \mathrm{tp}_{\boldsymbol{\Delta}_{n+1}}(\langle b_0, \ldots, b_{k-1}, y\rangle, \emptyset, M),$$

we fix $w = \{\ell < k : d(h(b_\ell), h(y)) \leq m(n)\}$.
E.g. if $w \neq \emptyset$ and $\ell(*) = \min(w)$, then the relevant properties of y are expressed in the balls

$$\{z \in M : d(h(z), h(b_\ell)) \leq m(n)\}$$

for $\ell \in w$ and

$$\{z \in M : d(h(z), h(y)) \le m(n)\},$$

all included in the ball

$$\{z \in M : d(h(z), h(b_{\ell(*)})) \le 3m(n)\}.$$

The case $w = \emptyset$ is simpler.

§5 Models with Distance

5.1 Discussion: We try here to explain the results of §2 as concerning a model with a distance function "weakly suitable" for the model and the connection to models with a distance function for the whole vocabulary which is suitable for the model. This is another variant of the distorted sums.

5.2 Context: Let τ be a fixed vocabulary.
1) Let K be the class of $\mathfrak{A} = (M, d)$, M a τ-model, d a distance on M (i.e. a two place symmetric function from $|M|$ to $\omega \cup \{\infty\}$, $d(x, x) = 0$, satisfying the triangular inequality) and for simplicity $d(x, y) = 0 \Leftrightarrow x = y$.
2) $K^{\rm sut} \subseteq K$ is the class of $(M, d) \in K$ such that d (which is a distance on M) is suitable for the model, i.e.

$\otimes_1$ $\langle a_0, \dots, a_{k-1}\rangle \in R^M, R \in \tau \;\Rightarrow\; \bigwedge_{\ell<m<k} d(a_\ell, a_m) \le 1$.

3) $K^{\rm sim} \subseteq K$ is the class of $(M, d) \in K$ which are simple, i.e

$\otimes_2$ $d(x, y) = \text{Min}\{n :$ there are $z_0, \dots, z_n$ such that $x = z_0, z_n = y$ and $\bigwedge_{\ell<n} \bigvee_{R\in\tau} (\exists \bar a \in R^M)[\{z_\ell, z_{\ell+1}\} \subseteq \text{Rang}\, \bar a]\}$.

4) $K^{\rm ws}_F$ is the class of $(M, d) \in K$ which are F-weakly suitable which means:

$\otimes_3$ for every m_i *if* $\bar a^i = \langle a^i_0, \dots, a^i_{n_i-1}\rangle$ for $i < k$, $a^i_\ell \in M$, $d(a^i_0, a^i_\ell) \le m_i$ and $i < j < k \Rightarrow d(a^i_0, a^j_0) > m_i + m_j + 1$ *then* the quantifier free type of $\bar a^0 {}^\frown \bar a^1 {}^\frown \dots {}^\frown \bar a^{k-1}$ is computed by F from the quantifier free types of $\bar a^0, \bar a^1, \dots, \bar a^{k-1}$ and of $\langle a^0_0, a^1_0, \dots, a^{k-1}_0\rangle$.

Note: we can strengthen the demands e.g. (for f as in 2.5)

(*) $d(a^i_0, a^i_\ell) \le r_i$, $d(a^i_0, a^j_0) > f_0(r_i) + f_0(r_j) + 1$ or at least $\neg(\exists x, y)[d(x, a_i) \le f_0(r_i) \wedge d(y, a_j) \le f_0(r_j) \wedge d(x, y) \le 1]$.

5) $K^{\rm as}$ is the family of $\mathfrak{B} = (M, d) \in K$ which are almost simple: "$d(x, y) \le 1$" is defined by quantifier free formula and $d(x, y) = \text{Min}\{n$: we can find $z_0, \dots, z_n, x = z_0, z_n = y$, $d(z_\ell, z_{\ell+1}) \le 1\}$.

5.3 Discussion: For $(M,d) \in K_F^{\rm ws}$ we want to "separate" the quantification to bounded ones and to distant ones. We can either note that it fits the context of §2 or repeat it.

5.4 Definition: 1) For $\mathfrak{B} = (M,d) \in K$, $x \in M$, $m < \omega$ let $N_m^+(x) = \{y \in M : d(y,x) \le m\}$.
2) We define "$\bar a$ is a $(\mathfrak{B}, r)$-component" and $\mathrm{bth}_r^n(\bar a, \mathfrak{B})$ as in Definition 2.8(1), and $\mathrm{BTH}_r^n(m,\tau)$ as in 2.8(2).
3) We define $\mathfrak{B}^2_{n,m}$ as expanding $\mathfrak{B}$ by the relation $R_{\mathbf{t}} = \{\bar a : \bar a$ is a $(\mathfrak{B}, r)$-component, $\mathbf{t} = \mathrm{bth}_r^n(\bar a, \mathfrak{B})\}$ for $\mathbf{t} \in \mathrm{BTH}_r^n(m,\tau)$.
4) We define "$\bar a$ is $(\mathfrak{B}, \bar r)$-sparse" and $\mathrm{uth}_{\bar r}^n(\bar a, \mathfrak{B})$", $\mathrm{UTH}_{\bar r}^n(m,\tau)$ as in 2.12.

5.5 Theorem: the parallel of 2.14.

References

[B] E.W. Beth, Observations metamathematiques sur les structures simplement ordonnes, *Collection de Logique Mathematique*, Serie A, Fasc 5, Paris–Louvain 1954, 29–35.

[CK] Chang and Keisler, **Model theory**, North Holland Publishing Co, 1973, 1977.

[Gf] H. Gaifman, On local and non local properties, Proc. Herbrand Symp. **Logic Colloquium, 81**, ed. J. Stern; Studies in Logic Found. of Math vol 107 North Holland (1982) 105–135.

[Gu] Y. Gurevich, Monadic second order theories, Chap XIII, **Model theoretic logics**, ed. Barwise and Feferman Springer Verlag (1985); 479–506.

[Lu] T. Luczak, notes.

[LcSh 435] T. Luczak and S. Shelah, Convergence in homogeneous random graphs, *Random Structures & Algorithms*, 1995, in press.

[ShSp 304] S. Shelah J. Spencer, On zero-one laws for random graph, *Journal of A.M.S* vol. 1 (1988) 97–115.

[Mo] A. Mostowski, On direct products of theories, *Journal of Symbolic Logic*, vol 17 (1952) 1–31.

[Mr] L. Marcus, Minimal models of theories of one function symbol, *Israel J. of Math* vol 18 (1974) 117–131.

[Sh 42] S. Shelah, The monadic theory of order, *Annals of Mathematics*, vol 102 (1975), 379–419.

[Sh F-120] S. Shelah, More on 0–1 laws, unpublished notes.

[Sh 467] S. Shelah, More on random finite models, in preparation.

[Sh 548] S. Shelah, Very weak zero one law for random graphs with order and random binary functions, *Random Structures & Algorithms*, submitted.

[Sp] J. Spencer, Survey/expository paper: zero one laws with variable probabilities, *Journal of Symbolic Logic* vol 58 (1993) 1–14.

[Sh 550] S. Shelah, 0–1 laws, in preparation.

[Sh 551] S. Shelah, In the random graph $G(n,p), p = n^{-a}$: if ψ has probability $0(n^{-\varepsilon})$ for every $\varepsilon > 0$ then it has probability $0(e^{-n^{\varepsilon}})$ for some $\varepsilon > 0$, in preparation.

[F] R. Fagin, Probabilities in finite models, *Journal of Symbolic Logic* vol 45 (1976) 129–141

[FV] S. Feferman and R. L. Vaught, The first-order properties of algebraic systems, *Fundamenta Mathematicae* vol 47 (1959) 57–103

[GKLT] Y.V. Glebskii, D.I. Kogan, M.I. Liagonkii and V.A. Talanov, Range and degree of reliability of formulas in restricted predicate calculus *Kibernetica* vol 5 (1969) 17-27; translation of *Cybernetics* vol 5 pp 142-154.