Henkin constructions of models with size continuum

John T. Baldwin*
Department of Mathematics
University of Illinois, Chicago

Michael C. Laskowski†
Department of Mathematics
University of Maryland

September 28, 2017

1 Introduction

In the novel *White Light* [Ruc80], Rudy Rucker proposes a metaphor for the continuum hypothesis. One can reach $\aleph_1$ by a laborious climb up the side of Mt. ON, pausing at ϵ_0. Or one can take Cantor's instantaneous elevator through the center of the mountain. In this paper, working in ZFC, we take Shelah's elevator, which is a bit slower. After countably many floors, each with finitely many rooms, we reach an object of cardinality $2^{\aleph_0}$. The underlying construction applies for finding atomic models, two-cardinal theorems, a collection of continuum many points that are *asymptotically similar* (a weak form of indiscernibility), and a coloring with a Borel square of size continuum.

In his seminal *Denumerable models of complete theories*, [Vau61], Vaught

*Partially supported by Simons grant MPS-SCG 418609
†Partially supported by NSF grant DMS-1308546. Both authors acknowledge the support of DMS 1362974 for visits to Rutgers.

introduced the notion of an atomic model[1]. He showed that if the isolated types were dense[2] in $S(T)$ then T has an atomic model. Interestingly, [HSS09] show that this central model theoretic theorem is not equivalent to any of the so-called 'big five' standard systems of reverse mathematics. Vaught further showed that a countable atomic model of a complete theory T could be elementarily embedded in every other model; that is, it is *prime*.

The construction of uncountable atomic models begins with Vaught's proof [Vau61] that if a countable atomic model has a proper atomic elementary extension then it has an atomic elementary extension of cardinality $\aleph_1$. He constructs a continuous, increasing sequence of ω_1 countable atomic models and, using the facts that unions of atomic models are atomic and elementarily equivalent countable atomic models are isomorphic, deduces the union of the chain is atomic. However the construction of atomic models in cardinals beyond $\aleph_1$ is a long standing problem. The study of atomic models of complete first order theories translates to the study of complete (decides every $L_{\omega_1,\omega}$-sentence) sentences of $L_{\omega_1,\omega}$ sentences. (See, e.g., Subsection 3.3 of this paper or Chapter 6 of [Bal09].).

Knight [Kni77] showed that construction could stop at $\aleph_1$; there is a first order theory with no atomic model of cardinality greater than $\aleph_1$. A series of works ([Kue78, LS93]) culminating in Hjorth [Hjo02] show that for each countable ordinal α there is a complete sentence of $L_{\omega_1,\omega}$ that has a model in $\aleph_\alpha$ but no larger. Thus, it is consistent that these sentences have no model in the continuum.

Given an atomic model M of cardinality $\aleph_1$ in a countable vocabulary, we describe simple sufficient conditions to construct an elementarily equivalent model N of cardinality $2^{\aleph_0}$, which is atomic and Borel. We modify Henkin's construction to build a complete diagram on a family of $2^{\aleph_0}$ variables. The traditional two steps in a Henkin construction, *completeness*, which ensures that each sentence is decided and *Henkin witnesses*, which ensures that each existential commitment is met, are supplemented by a crucial *splitting* stage which guarantees the final model has the cardinality of the continuum.

[1]Recall that a formula $\varphi(\overline{w})$, where $\lg(\overline{w}) = n$, is *complete* for T if for every formula $\psi(\overline{w})$, $\varphi(\overline{w})$ *decides* $\psi(\overline{w})$ in T. I.e. $T \vdash \forall\overline{w}[\varphi(\overline{w}) \rightarrow \psi(\overline{w})]$ or $T \vdash \forall\overline{w}[\varphi(\overline{w}) \rightarrow \neg\psi(\overline{w})]$. A model M is *atomic* if every finite tuple from A satisfies a complete formula. Here, atomic means φ is an atom in the Boolean algebra $F_n(T)$ and has nothing to do with the quantifier rank of the formula φ.

[2]For every formula $\varphi(\overline{x})$ consistent with T there is a complete formula $\psi(\overline{x})$ such that $T \vdash \forall\overline{x}[\psi(\overline{x}) \rightarrow \varphi(\overline{x})]$.

This method generalizes Shelah's construction of a kind of 'tree indiscernibility', which we call 'asymptotic similarity' to give a unified treatment of results in several areas of model theory. While we stressed atomic models in the first two paragraphs, the method applies as well to transfer cardinals in which a type is omitted and for two cardinal transfers.

We begin by describing the general method in the first five sections. Section 2 is an overview of both the classical Henkin construction and hints at the new construction. Section 3 lists a number of desirable properties we might wish the final model satisfied. Section 4 gives considerably more detail. There we define finite maximal antichains (fmacs) A of $2^{<\omega}$, A-commitments, and generating sequences. Theorem 5.4 of Section 5 is the main result of the paper.

The second half of the paper discusses applications of this technique. Most of the results are known, but Theorem 6.3.4 is new. Our first application in Subsection 6.1 constructs highly controlled models of theories with *trivial definable closure*, which is a notion studied by Ackerman, Freer, and Patel in [AFP16]. In Subsection 6.2 we introduce the notion of a *sufficient pregeometry* and prove, e.g., if M is uncountable and atomic and (M, cl) is a sufficient pregeometry, then there is an atomic model N of size continuum elementarily equivalent to M. This result immediately entails the new theorem that a pseudominimal theory has an atomic model of size continuum. In Subsection 6.4 we show that old results of Hrushovski and Shelah from [HS91] fit nicely into our rubric. In particular, if a superstable theory T has an atomic model of size $\aleph_1$, it has an atomic model of size $\beth_1$ (i.e. the continuum).

Section 7 is devoted to streamlining our method under the additional assumption that the theory T has Skolem functions. In Subsection 7.1 we show that Shelah's celebrated two-cardinal transfer theorem $(\aleph_\omega, \aleph_0) \to (2^{\aleph_0}, \aleph_0)$ from [She75b, She76] fits this framework. In Subsection 7.2, we discuss results of Shelah from [She99] that describe a cardinal $\lambda_{\omega_1}(\aleph_0)$ that is large enough so that any structure M of at least this size can witness arbitrarily long splittings. As one application, we expound Shelah's proof of the consistency with $ZFC + 2^{\aleph_0} > \aleph_{\omega_1}$ of the statement: 'A sentence of $L_{\omega_1,\omega}$ that has a model in $\aleph_{\omega_1}$ has one in the continuum.'

This analysis also connects with the philosophical discussion of the nature of mathematical explanation. Hafner and Mancosu [HM05] criticized the Resnik and Kushner [RK87] assertion that Henkin's proof [Hen49] of the completeness theorem for first order logic and type theory is explanatory.

They asked 'what the explanatory features of this proof are supposed to consist of?'. By its explicit connections with the deductive system Henkin's original proof was more explanatory of first order completeness than Gödel's reduction to propositional logic [Bal17]. This paper broadens that debate by noting that the Henkin construction extends from a transfer from a syntactic hypothesis to a semantic conclusion to a transformation from one model to another. That is, Henkin's essential contribution is to explain the ingredients to construct a model. So the significance of the method is seen in a larger context than the original proof.

2 General strategy

We suppose throughout that we are working with a countable language L with equality. Our objective will be to describe techniques, which are highly analogous to a Henkin construction of a countable model, for constructing a model M of size continuum.

Classically, the key notion is that of a *Henkin set* of formulas, whose definition is rather tedious, but provides the bridge between proof systems and structures. In their proofs of the completeness theorem both Henkin and Gödel worked in a framework in which equality was just another relation symbol. And each added an addendum that the proof transferred to the situation where equality was required to be interpreted as identity. This weakened (e.g. Henkin's) conclusion that the model constructed for a vocabulary of size κ had cardinal κ to 'at most κ' or Gödel version for countable languages allowed finite models. Because in our inductive construction we will introduce distinct variables that are later forced to be equal, we assume predicate logic includes the equality axioms, so *all witness sets will satisfy the usual equality axioms*. We are not giving a proof of the completeness theorem but transferring the existence of a model with specified properties to a model with the same properties but having cardinality $2^{\aleph_0}$.

Henkin's most fundamental innovation (e.g., [Bal17]) was to replace the Skolem functions in Gödel's proof by carefully described constants. This allowed the transformation from Gödel's universal vocabulary with relation symbols of all arities to a vocabulary tailored for the topic at hand.

Definition 2.1. Let L be any countable language. Let Z be a distinguished set of indexed variable symbols. After Henkin, Z was viewed as a countably infinite set of constant symbols. Here we treat the witnesses as variables so as to encode restrictions on the relations among variables introduced at different levels as transparent validities.

For any L-formula φ with at most k free variables and for any set of variables V, we introduce the notion of a *V-instantiated formula*. For any $(v_1, \ldots, v_k) \in V^k$, let $\varphi(v_1, \ldots, v_k)$ be the result of substituting the variable symbol v_j for the jth free variable for each j. We call $\varphi(v_1, \ldots, v_k)$ a *V-instantiated formula*; $Fm(V)$ denotes the set of all formulas obtained by this procedure.

A witnessed Henkin set is a subset $\mathcal{H} \subseteq Fm(Z)$ such that:

- **Satisfiable:** If $\varphi(z_1, \ldots, z_k) \in \mathcal{H}$, then there is some L-structure N and $(a_1, \ldots, a_k) \in N^k$ such that $N \models \varphi(a_1, \ldots, a_k)$.
- **Completeness:** For every $\varphi \in Fm(Z)$, exactly one of $\varphi, \neg\varphi \in \mathcal{H}$; and
- **Henkin witnesses:** If $\exists w \varphi \in Fm(Z)$, then either $\neg\exists w \varphi(w) \in \mathcal{H}$ or $\varphi(z^*) \in \mathcal{H}$ for some $z^* \in Z$.

It is routine to see that for any witnessed Henkin set $\mathcal{H} \subseteq Fm(Z)$, the binary relation $z \sim z'$ iff $(z = z') \in \mathcal{H}$ is an equivalence relation. As notation, for each $z \in Z$, let $[z]$ denote the image of z under the canonical projection $\pi : Z \to Z/\sim$. The following proposition is proved by a routine induction on the complexity of formulas; the 'Henkin witnesses' clause is precisely what is needed to allow quantifiers to be interpreted correctly.

Proposition 2.2. *If $\mathcal{H} \subseteq Fm(Z)$ is a witnessed Henkin set, then there is a unique L-structure M with universe $Z/\sim$ that satisfies*

$$M \models \varphi([z_1], \ldots, [z_k]) \qquad \Longleftrightarrow \qquad \varphi(z_1, \ldots, z_k) \in \mathcal{H}.$$

In particular, the relation $\sim$ induced by the equality symbol in $\mathcal{H}$ is a congruence on Z.

Moreover, if T is any L-theory and every $\varphi(z_1, \ldots, z_k) \in \mathcal{H}$ is satisfied by some model N of T (i.e., $N \models \varphi(a_1, \ldots, a_k)$ for some $(a_1, \ldots, a_k) \in N^k$), then M is a model of T.

Note that the whole of the discussion so far does not depend on the size of Z! In the classical construction of a Henkin set, Z is countably infinite, and $\mathcal{H}$ is generated by an ω sequence of formulas $\langle \varphi_n(\bar{z}_n) : n \in \omega \rangle$, where, for each n, $\bar{z}_n$ is a subsequence of $\bar{z}_{n+1}$ and $\varphi_{n+1}(\bar{z}_{n+1}) \vdash \varphi_n(\bar{z}_n)$. **In particular,**

at each finite stage and for each finite $\bar{z} \in Z^k$ only 'finitely much information' about $\mathcal{H}$ is determined.

In analogy with this construction, we want to create a template which can be customized to create a model of size $2^{\aleph_0}$ with desirable properties. We begin with an indexed set Z of variable symbols of cardinality $2^{\aleph_0}$, which are subdivided as

$$Z = \bigcup \{Z_s \colon s \text{ a non-empty finite subset of } 2^\omega\}$$

where each Z_s is countably infinite and $Z_t \subseteq Z_s$ whenever $t \subseteq s$.

We will construct a witnessed Henkin set $\mathcal{H} \subseteq Fm(Z)$ in ω steps. Our subdivision of Z gives rise to sets $Fm(Z_s)$ of instantiated formulas, whose intersection with $\mathcal{H}$ yields a family $\{\mathcal{H}(Z_s) \colon s$ a non-empty finite subset of $2^\omega\}$ of *countable* witnessed Henkin sets. The restrictions of the congruence $\sim$ on Z naturally induce congruences on each Z_s. Thus, exactly as in the classical case outlined above, each of the Henkin sets $\mathcal{H}(Z_s)$ gives rise to a canonical countable L-structure $M(s)$ with universe $Z_s/\sim$. Our construction will ensure that $M(t)$ is an elementary submodel of $M(s)$ whenever $t \subseteq s$.

Additionally, the entire Henkin set $\mathcal{H}(Z)$ determines a canonical L-structure M with universe $Z/\sim$. Since any finite tuple $\bar{z}$ from Z is contained in some Z_s, M can be identified with

$$M = \bigcup \{M(s) : s \text{ a non-empty finite subset of } 2^\omega\}$$

In particular, any 'finitary information' about M will be inherited from the directed family $\{M(s)\}$ of countable models. As examples,

- $M(s) \preceq M$ for each finite $s \subseteq 2^\omega$, hence for any T, $M \models T$ if and only if some (equivalently, every) $M(s) \models T$;
- For Δ any partial type, M omits Δ if and only if every $M(s)$ omits Δ; so
- M is atomic (Subsection 3.2) if and only if every $M(s)$ is atomic.

Obviously, if we want to conclude that M has size $2^{\aleph_0}$, we need some additional mechanism to ensure the *construction is non-degenerate*. In particular, as each $M(s)$ is countable, it would be very unfortunate if $M(s) = M(t)$ for all finite subsets s, t!

To ensure this, we now introduce the actual set of variable symbols used in the construction. We will write $Z = X \cup Y$, where, X is indexed as $\{x_\eta \colon \eta \in 2^\omega\}$ or sometimes we must doubly index X as $\{x_{\eta,i} \colon \eta \in 2^\omega, i \in \omega\}$.

The intent is that the elements of X are 'independent' in some sense; but at a minimum, we will require that for distinct η, η', $x_\eta \neq x_{\eta'} \in \mathcal{H}$[3]. This will be enough to guarantee that the model M we produce from $\mathcal{H}$ will have power continuum. The Y-symbols are indexed as $\{y_{s,i} \colon s$ a non-empty finite subset of 2^ω, $i \in \omega\}$ and should be interpreted as collectively being 'material needed to close X into a model.' For each non-empty finite subset s of 2^ω, put $X_s := \{x_\eta \colon \eta \in s\}$ (or $\{x_{\eta,i}, \eta \in s, i \in \omega\}$ in the doubly-indexed case); put $Y_s := \{y_{t,i} \colon t \subseteq s, i \in \omega\}$, and $Z_s := X_s \cup Y_s$. Visibly, each Z_s is countable and $Z_t \subseteq Z_s$ whenever $t \subseteq s$.

As examples, consider the models $M_{\{\eta\}}$, $M_{\{\eta'\}}$ and M_s, where $s = \{\eta, \eta'\}$. Each of these is a countable, elementary substructure of M. Thus, in particular, for every constant symbol $c \in L$, there will be natural numbers i, j such that the Z-instantiated formulas $y_{\eta,i} = c$ and $y_{\eta',j} = c$ are both in $\mathcal{H}$. Consequently, $y_{\eta,i} = y_{\eta',j}$ will also be in $\mathcal{H}$, so $y_{\eta,i} \sim y_{\eta',j}$. That is, these two variable symbols are identified in both M and $M(s)$.

For $s = \{\eta, \eta'\}$, the variables for $M(s)$ are the union of the variables of $M(\eta)$, $M(\eta')$ and $\{y_{s,i}\}$ for $i < \omega$. The additional variables $\{y_{s,i}\}$ will close $M(s)$ to be a model. For example, if we are constructing a group, then for some $i \in \omega$, $\mathcal{H}$ would include the Z-instantiated formula $x_\eta + x_{\eta'} = y_{\{\eta,\eta'\},i}$.

3 Desirable properties of models

As we are working in a countable language, the existence of structures, or even models of a consistent first order theory, of size continuum is not surprising. Our aim is to identify other desirable properties of models that do not so obviously have uncountable models but that can be dovetailed with our construction of a witnessed Henkin set. Here, we describe some such properties, and the next section will outline sufficient conditions for a generating sequence and hence a witnessed Henkin set to admit these properties.

[3] Or $x_{\eta,0} \neq x_{\eta',0} \in \mathcal{H}$ in the doubly indexed case.

3.1 Modeling T and omitting types

We list here the goals of certain conditions on a construction that will guarantee it yields a model of a given theory T that has the properties we are after. In Definition 5.3, we specify how these goals are met in our situation.

Modeling T: As L-sentences are themselves L-formulas, if we require every $\varphi(\bar{z}) \in \mathcal{H}$ to be satisfiable in some model of T, then the **Completeness** condition, each L-formula φ or its negation is in $\mathcal{H}$, on a witnessed Henkin set will ensure that the canonical model M built from $\mathcal{H}$ is a model of T.

Omitting Δ: If we want M to omit a single partial type Δ we need to require that for any $\bar{z} \in Z^k$, there is some $\delta \in \Delta$ with $\neg\varphi(\bar{z}) \in \mathcal{H}$. So, if $\mathcal{H}$ is going to be produced in ω steps, we need to ensure that every $\bar{z} \in Z^k$ is 'handled' along the way. Note that, in general, a condition such as 'every $\varphi \in \mathcal{H}$ is realized in some model that omits Δ' might not be sufficient to guarantee that M omits Δ.

Omitting $\{\Delta_m : m \in \omega\}$: Similarly, if we are given a countable set $\{\Delta_m\}$ of partial types, in order to ensure that M omits each Δ_m, we need to ensure that for each pair $(\bar{z}, m)$, there is a $\delta \in \Delta_m$ for which we enforce that $\neg\delta(\bar{z}) \in \mathcal{H}$.

3.2 Atomic models and complete formulas

For a complete theory T, an L-formula $\varphi(\overline{x})$ is *complete with respect to* T if:

- $T \models \exists\overline{x}\varphi(\overline{x})$ and;
- for every L-formula $\delta(\overline{x})$, φ *decides* δ,
  - either $T \models \forall\overline{x}(\varphi(\overline{x}) \to \delta(\overline{x}))$;
  - or $T \models \forall\overline{x}(\varphi(\overline{x}) \to \neg\delta(\overline{x}))$.

Equivalently, $\varphi(\overline{x})$ is complete with respect to T if and only if there is a unique complete type extending $\varphi(\overline{x})$.

A model M of T is *atomic* if, for every $n \geq 1$, every tuple $\bar{a} \in M^n$ realizes a complete formula with respect to T. Not every countable theory T admits an atomic model, but Vaught proved that any two countable, atomic models are isomorphic. It is easy to see that any elementary submodel of an atomic model is atomic, but the Upward Löwenheim-Skolem theorem can fail badly

– Hjorth [Hjo07] proved that for any $\alpha < \omega_1$, there are complete theories T_α that have atomic models of size $\aleph_\alpha$, but no larger. As it is consistent with ZFC for the continuum to be arbitrarily large in the $\aleph$-hierarchy, we know that we cannot hope to construct an atomic model of size continuum for any of these theories T_α. So we must impose some additional hypotheses on T for it to have an atomic model in the continuum.

3.3 $L_{\omega_1,\omega}$-sentences, omitting types, atomic models

We will see that in many cases, the Henkin method will provide sufficient conditions for building a model of size continuum that is atomic, or, in other cases, omits a given countable family of types. This dual consequence stems from a fundamental link, discovered independently by Chang and Lopez-Escobar, between sentences[4] Φ of $L_{\omega_1,\omega}$ and the omitting of types, which Shelah extended to atomic models.

> Given any sentence Φ' of $L_{\omega_1,\omega}$ there is a countable language $L' \supseteq L$, a first-order L'-theory T, and a partial L'-type $\Delta(w)$ such that the class of models of Φ' is precisely the class of L-reducts of models of T that omit $\Delta(w)$.

To see the idea suppose a subformula $\Phi(\overline{w})$ of the sentence Φ' is a countable conjunction of formulas $\varphi_i(\overline{w})$. Add a new predicate symbol $R_\Phi(\overline{w})$. Let T assert for each i, $\forall\overline{w}[R_\Phi(\overline{w}) \rightarrow \varphi_i(\overline{w})]$ and let $\Delta(\overline{w})$ be the type $\{\neg R_\Phi(\overline{w})\} \cup \{\varphi_i(\overline{w}) : i < \omega\}$. Now a model M satisfies $\Phi(\overline{w}) \leftrightarrow R_\Phi(\overline{w})$ if and only if M omits $\Delta(w)$. Now hire a secretary who translates the inductive structure of arbitrary sentence Φ' into an iteration of extensions of this sort.

To make the connection with atomic models, we need some further terminology.

Definition 3.3.1. An $L_{\omega_1,\omega}$-sentence Φ is *complete* if it has a model and if it decides every $L_{\omega_1,\omega}$-sentence Ψ. An L-structure M is *small* if it realizes only countably many distinct $L_{\infty,\omega}$-types over the empty set.

Recall that each countable model M (in a countable vocabulary) has a Scott sentence, an $L_{\omega_1,\omega}$-sentence Φ_M, whose only model is M. By the Löwenheim Skolem theorem Φ_M is complete. Examining the proof of Scott's

[4]Recall that the logic $L_{\kappa,\omega}$ allows conjunctions of length less than κ but only finite quantifications; $L_{\infty,\omega} = \cup_\kappa L_{\kappa,\omega}$.

theorem ([Kei71]) one sees several equivalent statements (see e.g., Chapter 6 of [Bal09]): an $L_{\omega_1,\omega}$-sentence Φ is complete if and only if Φ is $\aleph_0$-categorical if and only if Φ is a Scott sentence of a countable L-structure. Similar arguments show that an L-structure M is small if and only if it satisfies a complete sentence Φ if and only if it has a countable $L_{\infty,\omega}$-elementary substructure if and only if it has a countable $L_{\omega_1,\omega}$-elementary substructure.

Shelah [She75a] observed:

Remark 3.3.2. *If Φ is a complete $L_{\omega_1,\omega}$-sentence, then there is a countable language $L' \supseteq L$ and an L'-structure M' such that the class of models of Φ is precisely the class of L-reducts of atomic models of $T = Th(M')$. Conversely, given any complete theory T in a countable language, there is a complete sentence Φ of $L_{\omega_1,\omega}$ whose models are precisely the atomic models of T.*

Proof. Let M be any countable model of Φ. For each $k \geq 1$, define an equivalence relation $\sim_k$ on M^k by $\bar{a} \sim_k \bar{b}$ if and only if they have the same $L_{\infty,\omega}$-type over the empty set. For each k and $\sim_k$-class E, add a new, k-ary predicate symbol R^k_E to L' and let M' be the natural expansion of M, i.e., $M' \models R^k_E(\bar{a})$ if and only if $\bar{a} \in E$. Let $T = Th(M')$.

Conversely, given a complete, first order theory T, for every n let $\Delta_n(\overline{x})$ be the partial type asserting the negation of every complete formula with respect to T. Let Φ be the $L_{\omega_1,\omega}$-sentence

$$\bigwedge T \wedge \bigwedge_n \forall \overline{x} \left(\neg \bigwedge \Delta_n(\overline{x}) \right)$$

The models of Φ are precisely the atomic models of T. The completeness of Φ follows from the uniqueness of countable, atomic models of T.

Because of these observations, the entire subfield of 'atomic model theory' can be considered to be a study of the classes of models of complete sentences of $L_{\omega_1,\omega}$. Shelah exploited this identification by studying atomic models to generalize Morley's categoricity theorem to $L_{\omega_1,\omega}$ in [She83a, She83b].

3.4 Borel structures

Following [MN13], we say that a structure M is *Borel* if there is a standard Borel space Z, a Borel subset $D \subseteq Z$, and a congruence $E \subseteq Z^2$ such that

1. E is a Borel subset of Z^2;

2. The universe of M is D/E; and
3. The pre-image of every subset of M^k defined by an atomic formula is a Borel subset of Z.

If the congruence is the identity, we say that M has an *injective presentation.*

In all cases we consider, the set Z of variable symbols can be presented as a standard Borel space. As we construct the witnessed Henkin set $\mathcal{H}$ (which yields the entire elementary diagram of Z) in ω steps, it will follow automatically that the associated model M is a Borel structure, where, moreover $D = Z$. Typically, however, our methods do not give an injective presentation of M. The one exception to this is in Section 6.1, where we exploit strong hypotheses (trivial definable closure) about the theory that yield an injective presentation. In that case, we additionally show that every definable subset of M^k is a finite Boolean combination of open sets.

3.5 Asymptotic similarity

Throughout his career, Saharon Shelah defined and reaped the benefits from a weakish notion of indiscernibility, that he used in many varied contexts, including two cardinal transfer theorems in [She75b, She76], obtaining perfect squares of colorings as in [She99], and constructing many models in small, superstable, non-$\aleph_0$-stable theories. Until now, this notion was unnamed; we give it a belated baptism as *asymptotic similarity.*

In order to describe this notion we fix some notation for dealing with sequences from 2^ω

Definition 3.5.1. Fix an integer ℓ.

- A k-tuple $(\eta_0, \ldots, \eta_{k-1})$ of distinct elements from 2^ω *splits by* ℓ if the restrictions $\{\eta_i\restriction_\ell : i < k\}$ to 2^ℓ are distinct.
- Two k-tuples $(\eta_0 \ldots, \eta_{k-1})$ and $(\tau_0, \ldots, \tau_{k-1})$ of distinct elements from 2^ω are *similar (mod ℓ)* if $(\eta_0, \ldots, \eta_{k-1})$ splits by ℓ and $\eta_i\restriction_\ell = \tau_i\restriction_\ell$ for each $i < k$.

Clearly, every k-tuple of distinct elements from 2^ω splits by some ℓ, and consequently splits by every $\ell' \geq \ell$; and similarity (mod ℓ) is an equivalence relation on the set of k-tuples from 2^ω that split by ℓ.

Definition 3.5.2. Fix an L-structure M. A subset of M, indexed by $\{a_\eta : \eta \in 2^\omega\}$, is *asymptotically similar* if, for every k-ary L-formula θ, there is an integer N_θ such that for every $\ell \geq N_\theta$,

$$M \models \theta(a_{\eta_0}, \ldots, a_{\eta_{k-1}}) \leftrightarrow \theta(a_{\tau_0}, \ldots, a_{\tau_{k-1}})$$

whenever $(\eta_0, \ldots, \eta_{k-1})$ and $(\tau_0, \ldots, \tau_{k-1})$ are similar (mod ℓ).

Remark 3.5.3. Although asymptotic similarity should be thought of as a type of indiscernibility, the indiscernibility is only formula by formula. For example, consider the structure $M = (2^\omega, U_a)_{a \in 2^{<\omega}}$, where each U_a is a unary predicate interpreted as the cone above a, i.e., $U_a(M) = \{\eta \in 2^\omega : a \lhd \eta\}$. Then, in M, the entire universe $\{\eta : \eta \in 2^\omega\}$ is asymptotically similar, despite the fact that no two elements have the same 1-type.

This notion of indiscernibles should not be confused with the 'tree-indexed indiscernibles' (which are indiscernible for all formulas in the vocabulary) in [KKS14] which arise from non-superstable theories and Theorem 3.6 of [She78].

4 Partitions of Z via finite antichains

A cursory inspection shows that the set 2^ω is involved in the indexing of elements from Z. We employ the standard topology placed on the space 2^ω to describe families of partitions of Z. As notation, for any $a \in 2^{<\omega}$, let $U_a = \{\eta \in 2^\omega : a \lhd \eta\}$ and $\mathcal{U} = \{U_a : a \in 2^{<\omega}\}$. The *standard topology* on 2^ω is the topology formed by positing that $\mathcal{U}$ is a base of open sets.

> **Throughout this paper, we will denote elements of $2^{<\omega}$ by lower case roman letters, $a, b, c, \ldots$, and we reserve lower case Greek letters $\eta, \nu, \ldots$ for elements of 2^ω.**

Note that if two elements $a, b \in 2^{<\omega}$ are *incomparable,* i.e., $a \not\trianglelefteq b$ and $b \not\trianglelefteq a$, then the sets U_a and U_b are disjoint. A *finite, maximal antichain*, abbreviated *fmac* is a finite set $A \subseteq 2^{<\omega}$ in which any two elements are incomparable, and every $b \in 2^{<\omega}$ is comparable to some $a \in A$. It is easily seen that if A is an fmac, then the sets $\{U_a : a \in A\}$ form a partition of 2^ω. As notation, let $\pi_A : 2^\omega \to A$ denote the projection map, i.e., $\pi_A(\eta)$ is the unique element of A lying below η. Curiously, the restriction that A is finite is crucial to obtain a partition of 2^ω. Indeed, if A is any *infinite* antichain, then as 2^ω is compact and each of the sets U_a are clopen, $\{U_a : a \in A\}$ cannot cover 2^ω. Paradigms of fmacs are the sets 2^n, consisting of all sequences of length n, but many other fmacs exist. Our constructions could be done using only the sets 2^n but at the cost of suppressing intermediate steps which are fmacs; it is more convenient to do various inductions in the general setting.

We now introduce a second system of variables. Given any fmac $A \subseteq 2^{<\omega}$, let Z_A be the following set of variable symbols that are *disjoint from* Z. The indexing on Z_A will parallel that for Z. In particular, Z_A is partitioned into $X_A \cup Y_A$, X_A is either indexed as $\{x_a : a \in A\}$ or doubly indexed as $\{x_{a,i} : a \in A, i \in \omega\}$, and $Y_A = \{y_{t,i} : t \subseteq A, i \in \omega\}$. For a subset $s \subseteq A$, the sets X_s and Y_s are defined analogously. Note that in the definition that follows, we build in both the **Satisfiable** condition, as well as a *'non-degeneracy'* condition that will imply that the Henkin model we construct has size continuum.

Definition 4.1. Let $A \subseteq 2^{<\omega}$ be any fmac. Define an A-*commitment* to be a Z_A-instantiated formula

$$\varphi(\overline{x}, \bar{y}), \text{ where } \overline{x} = \langle x_a : a \in A\rangle \text{ and } \bar{y} \subseteq Y_A$$

that is satisfiable in some L-structure and with the additional property that for each $a, a' \in A$, $\varphi \vdash x_a \neq x_{a'}$ (or $x_{a,0} \neq x_{a',0}$ when X_A is doubly indexed).

To understand the relevance of an A-commitment to a Henkin set $\mathcal{H}$ we are constructing, we need the notion of a *lifting* $h^*\colon A \to 2^\omega$ of the fmac A to 2^ω, which is any (necessarily injective) mapping satisfying $a \triangleleft h^*(a)$ for every $a \in A$. Note that any lifting h^* naturally induces an injection, which we also dub h^*,

$$h^*\colon Fm(Z_A) \to Fm(Z)$$

given by replacing each x_a by $x_{h^*(a)}$ and replacing each $y_{s,i}$ by $y_{h^*(s),i}$, where $h^*(s) = \{h(a) : a \in s\}$.

Our intent is that if, at some stage of our construction of $\mathcal{H}$ we include the A-commitment φ, we commit ourselves to eventually making

$$\{h^*(\varphi)\colon \text{ all liftings } h^* : A \to 2^\omega\}$$

a subset of $\mathcal{H}$. More precisely, we define:

Notation 4.2. A *commitment* is a pair (A, φ), where A is an fmac and φ is an A-commitment. Each construction will choose a particular set of A-commitments (for enough A) to determine the diagram of Z.

Given two fmacs A and B, we say that B *covers* A, written $A \leq B$, if, for every $a \in A$ there is at least one $b \in B$ such that $a \trianglelefteq b$. For example, if $n \leq m$, then 2^m is a cover of 2^n.

If $A \leq B$, then a *lifting to* B is a (necessarily injective) map $h \colon A \to B$ satisfying $a \trianglelefteq h(a)$ for each $a \in A$. Note that if $A \leq B$, then any lifting $h^* \colon A \to 2^\omega$ factors through B. That is, given any lifting $h^* \colon A \to 2^\omega$, define $h_B \colon A \to B$ by $h_B(a) = \pi_B(h^*(a))$ (where π_B is the natural projection from 2^ω onto B). Any such h_B is a lifting to B, and there is a natural lifting $h' \colon B \to 2^\omega$ satisfying $h^* = h' \circ h_B$.

With this in mind, we partially order the set of commitments by:

> $(A, \varphi) \leq (B, \psi)$ if and only if B covers A and[5] $\psi \vdash h(\varphi)$ for every lifting $h : A \to B$.

We say (B, ψ) *extends* (A, φ) when $(A, \varphi) \leq (B, \psi)$. Because of our comments about compositions of liftings, it is evident that whenever (B, ψ) extends (A, φ), what ψ commits us to about the $\mathcal{H}$ we will construct is consistent with, and typically extends what φ commits us to about $\mathcal{H}$. Thus, if we have an ω-sequence $\overline{A} = \langle (A_n, \varphi_n) \colon n \in \omega \rangle$ of commitments such that $(A_n, \varphi_n) \leq (A_{n+1}, \varphi_{n+1})$ for each n, then let

> $\mathcal{D}_{\overline{A}} := \{Z$-instantiated formulas $\theta(\bar{z})$: for some n (equivalently, for all sufficiently large n) there is some lifting $h^* : A_n \to 2^\omega$ such that $h^*(\varphi_n) \vdash \theta(\bar{z})\}$.

Visibly, any such set $D_{\overline{A}}$ is closed under logical consequence. It is natural to ask for sufficient conditions for a sequence of commitments to determine a witnessed Henkin set. More formally:

Definition 4.3. A *generating sequence* is a $\leq$-increasing ω-sequence $\overline{A} = \langle (A_n, \varphi_n) : n \in \omega \rangle$ of commitments such that $\mathcal{D}_{\overline{A}}$ is a witnessed Henkin set.

By coupling the discussion in this section with Proposition 2.2, we see that if $\overline{A} = \langle (A_n, \varphi_n) : n \in \omega \rangle$ is a generating sequence, then $\mathcal{D}_{\overline{A}}$ uniquely describes a model M of size $2^{\aleph_0}$.

[5]The $\vdash$ means that $(\forall \bar{z})[\psi \to h(\varphi)]$, where $\bar{z}$ lists the free variables of the formula, is a theorem of the predicate calculus; it is to state this clearly that we work with variables rather than constants.

5 Sufficient conditions for producing Henkin models of size continuum

We now describe the machinery for constructing a generating sequence. Even though our construction is in ZFC, cognoscenti will recognize the affinity of our nomenclature with that of forcing. We begin by discussing properties of partially ordered sets $(\mathbb{P}, \leq)$ of commitments. Note that the 'classical Henkin constraints', laid down in the definition of a witnessed Henkin set, of **Completeness** and **Henkin witnesses** can be phrased in terms of showing that certain subsets of $\mathbb{P}$ are dense and open[6] in $(\mathbb{P}, \leq)$. Additionally, the **Satisfiable** condition is built into the definition of an A-commitment. The additional density condition we need to allow us to simultaneously construct the family $\{M(s)\colon s \text{ a non-empty finite subset of } 2^\omega\}$ of countable models is **Splitting**.

Definition 5.1. Given any fmac A and any $a \in A$, the *splitting of A at a* is the fmac $A^{*a} = A \setminus \{a\} \cup \{a\hat{\ }0, a\hat{\ }1\}$. Clearly, A^{*a} covers A, and there are two liftings $h_0, h_1 : A \to A^{*a}$, distinguished by $h_i(a) = a\hat{\ }i$ for $i = 0, 1$. Thus, by the definition of extension, if an A^{*a}-commitment φ^* *extends* an A-commitment φ then $\varphi^* \vdash h_0(\varphi) \wedge h_1(\varphi) \wedge x_{a\hat{\ }0} \neq x_{a\hat{\ }1}$.

It is an easy exercise to verify that whenever an fmac B covers A, then B can be obtained by a sequence of splittings at points. Indeed, the fmac 2^{n+1} can be obtained from 2^n by a sequence of 2^n splittings, one at each $a \in 2^n$. The following notation will be used to ensure that appropriate Henkin witnesses are put into a Henkin set.

Definition 5.2. Given any fmac A and any finite tuple $\bar{z}$ from Z_A, let $t(\bar{z})$ denote the smallest subset of A for which $\bar{z} \in Z_{t(\bar{z})}$.

Unpacking the definitions, $t(\bar{z})$ is the smallest subset of A that satisfies (1) If $x_a \in \bar{z}$, then $a \in t(\bar{z})$; and (2) if $y_{s,i} \in \bar{z}$, then $s \subseteq t(\bar{z})$.

Definition 5.3. A set $(\mathbb{P}, \leq)$ of commitments, ordered by extension, is *sufficiently dense* if, for every fmac A and every A-commitment $\varphi \in \mathbb{P}$ we have:

[6]We use Shelah's convention that 'more information' puts you 'higher up' in $(\mathbb{P}, \leq)$. Thus, X is *dense* in $(\mathbb{P}, \leq)$ if for every $q \in \mathbb{P}$, there is an $x \in X$ with $p \leq x$. X is *open* if $q \in X$ whenever $q \geq x$ for some $x \in X$.

- **Completeness:** For every Z_A-formula ψ, there is an A-commitment $\varphi^* \in \mathbb{P}$ extending φ that decides ψ. By 'decides', we mean **either** $\varphi^* \vdash \psi$ **or** $\varphi^* \vdash \neg\psi$;
- **Henkin Witnesses:** For every $\theta(u, \overline{w})$ and every $\bar{z} \in (Z_A)^{\lg(\overline{w})}$, there is an A-commitment $\varphi^* \in \mathbb{P}$ extending φ such that **either** $\varphi^* \vdash \forall u \neg\theta(u, \bar{z})$ **or** $\varphi^* \vdash \theta(z^*, \bar{z})$ for some $z^* \in Z_{t(\bar{z})}$.
- **Splitting:** For every $a \in A$ there is an A^{*a}-commitment $\varphi^* \in \mathbb{P}$ extending φ. [In particular, $\varphi^* \vdash h_0(\varphi) \wedge h_1(\varphi) \wedge x_{a\hat{\ }0} \neq x_{a\hat{\ }1}$.]

Before stating the main theorem, we specify in our context the properties ensuring the goals laid out at the beginning of Section 3.1. They may or may not hold of a particular $(\mathbb{P}, \leq)$:

- **Modeling T:** Given a theory T, if a condition $(A, \varphi) \in \mathbb{P}$, then φ is satisfiable in some model of T.
- **Omitting a type $\Delta(\overline{w})$:** For every A-commitment $\varphi \in \mathbb{P}$ and every $\bar{z}$ from Z_A, there is a some $\delta \in \Delta$ and an A-commitment φ^* extending φ with $\varphi^* \vdash \neg\delta(\bar{z})$.
- **Atomic model:** Given a complete theory T, whenever $(A, \varphi) \in \mathbb{P}$, φ is a complete formula (in its free variables) with respect to T.

Theorem 5.4. *Let T be any theory in a countable language. If there is a sufficiently dense, partially ordered set $(\mathbb{P}, \leq)$ of commitments that are each satisfied in a model of T, then there is a Borel model M of T of size continuum with an asymptotically similar subset $\{a_\eta : \eta \in 2^\omega\}$. Moreover:*

1. *If $\{\Delta_m : m \in \omega\}$ is a countable set of partial types*[7] *and if $(\mathbb{P}, \leq)$ satisfies* **Omitting** Δ_m *for each m, then such an M can be chosen to omit each Δ_m; and*
2. *If T is complete and if $(\mathbb{P}, \leq)$ satisfies the* **Atomic model** *condition, then such an M can be chosen to be an atomic model of T.*

[7] So the Δ_m each exemplify a $\Delta(\overline{w})$ in Definition 5.3.

Proof. Fix a distinguished set $Z = X \cup Y$ of variable symbols, for definiteness[8], say $X = \{x_{\eta,i} : \eta \in 2^\omega, i \in \omega\}$ and $Y = \{y_{t,i} : t$ a finite subset of 2^ω and $i \in \omega\}$.

The following notation will be helpful. For a fixed $\ell \in \omega$, consider the 'standard fmac' 2^ℓ. In order to consider only finitely many Y-variables at each stage, we distinguish a sufficiently large, finite subset of symbols in $Z_{(2^\ell)}$.

Let

$$W_\ell := \{x_{a,i} : a \in 2^\ell, i < \ell\} \cup \{y_{t,i} : t \subseteq 2^\ell, i < \ell\}.$$

Note that W_ℓ is a finite subset of $Z_{(2^\ell)}$ and, whenever $\ell \leq m$, $h(W_\ell) \subseteq W_m$ for every lifting $h : 2^\ell \to 2^m$. We will construct a generating sequence $\overline{A} = \langle (A_n, \varphi_n) : n \in \omega \rangle$ from $\mathbb{P}$ in ω steps. We will dovetail these extensions to obtain the following goals:

(i) All but finitely many of the 'standard fmacs' 2^ℓ will appear as A_n's in our generating sequence;

(ii) To obtain asymptotic similarity, for every formula $\psi(\overline{w})$ there is a number N_ψ such that for all $\ell \geq N_\psi$ there is an n such that $A_n = 2^\ell$ and, for every $\bar{z}$ from W_ℓ, φ_n decides $\psi(\bar{z})$;

(iii) To show that each of the countable models $M(s) \preceq M$, we require that for every formula $\theta(u, \overline{w})$ there is a number N_θ such that for all $\ell \geq N_\theta$ there is an n such that $A_n = 2^\ell$ and, for every $\bar{z}$ from W_ℓ, either $\varphi_n \vdash \neg\exists u \theta(u, \bar{z})$ or $\varphi_n \vdash \theta(y_{t(\bar{z}),i^*}, \bar{z})$ for some $i^* \in \omega$ (recall Definition 5.2);

(iv) Depending on whether we are verifying 1) or 2) there are two further conditions.

  (a) For each partial type $\Delta_m(\overline{w})$ we are asked to omit, there will be some $N(m)$ such that for every $\ell \geq N(m)$, there is an n such that $A_n = 2^\ell$ and, for every $\bar{z}$ from W_ℓ (of length $\lg(\overline{w})$) there is $\delta \in \Delta_m$ such that $\varphi_n \vdash \neg\delta(\bar{z})$;

  (b) Finally, if we are asked to produce an atomic model, we require either that every element of $\mathbb{P}$ be a complete formula, or that for all but finitely many ℓ, there is an n such that $A_n = 2^\ell$ and, for every $\bar{z}$ from W_ℓ, φ_n entails some complete formula $\eta(\bar{z})$.

[8]The a_η will be the interpretations of the $x_{\eta,0}$ for $\eta \in 2^\omega$.

How can we construct such a generating sequence? We systematically extend an arbitrary fmac to an A of the form 2^ℓ that satisfies the appropriate condition. Satisfying (i) is straightforward. Indeed, given any $(A, \varphi) \in \mathbb{P}$, choose any ℓ such that 2^ℓ covers A. Then, as noted in the discussion above, 2^ℓ can be obtained from A by a sequence of splittings at points. So, it follows from a finite number of applications of **Splitting** that there is some sequence $\langle (B_0, \varphi_0), \dots (B_n, \varphi_n) \rangle$ from $\mathbb{P}$ with $B_0 = A$, $B_n = 2^\ell$, and (B_{i+1}, φ_{i+1}) extends (B_i, φ_i) for each $i < n$.

To handle (ii) and (iii), fix an enumeration of L-formulas $\{\psi_i(\overline{w}) : i < \omega\}$ and $\{\theta_i(u, \overline{w}) : i < \omega\}$. For (ii), observe that as each W_ℓ is finite, there are only finitely many instantiations $\psi_i(\bar{z})$ with both $i < \ell$ and $\bar{z}$ from W_ℓ. Thus, using the **Completeness** condition on $(\mathbb{P}, \leq)$ finitely many times, given any (A_n, φ_n) with $A_n = 2^\ell$, there is an extension $(2^\ell, \varphi_{n+1}) \geq (2^\ell, \varphi_n)$ in which φ_{n+1} decides every $\psi_i(\bar{z})$ with $i < \ell$ and $\bar{z}$ from W_ℓ.

Similar remarks concern clause (iii). Here, the formulas $\{\theta_i(u, \bar{z}) : i < \ell\}$ apply, where we use the **Henkin witnesses** condition finitely often. Continuing, again because W_ℓ is finite, we can use **Omitting** Δ_m or **Atomic** to further extend to some $(2^\ell, \varphi_j) \in \mathbb{P}$ with $j \geq n$ that satisfy iv(a) or iv(b).

Now, once we have handled all of our requirements for the fmac 2^ℓ, note that $2^{\ell+1}$ covers 2^ℓ, so by finitely many applications of **Splitting** we get an extension (A_{n+1}, φ_{n+1}) with $A_{n+1} = 2^{\ell+1}$, thus completing (i) for the next step. We repeat the discussion above, but now with the larger $A_{n+1} = B_{\ell+1}$ and a larger (finite) set of formulas $\psi_i(\overline{w}) \in W_{\ell+1}$ and $\theta_i(u, \overline{w})$, for $i < (\ell+1)$.

Continuing this for ω steps gives us a generating sequence $\overline{A} = \langle (A_n, \varphi_n) : n \in \omega \rangle$ from $\mathbb{P}$. As cofinally many of the A_n's are 2^ℓ for increasing ℓ's, it follows that $\mathcal{D}_{\overline{A}}$ describes a complete type in the variables Z. The non-degeneracy condition in the definition of a commitment will imply that $\{x_{\eta,0} : \eta \in 2^\omega\}$ are pairwise distinct. Also, by (ii), this set is easily seen to be asymptotically similar.

In the construction above, for any witnessed existential formula, for all but finitely many ℓ, a witness was placed in $Z_{(2^\ell)}$. Thus, one can check that if s is a finite subset of 2^ω, then $M(s) := \{[z] : z \in Z_s\}$ is a countable model and $M(s) \preceq M$. As well, Clause iv(a) will imply that $M(s)$ omits each Δ_m, and, in the atomic case, iv(b) ensures that $M(s)$ is atomic. As noted in Section 2, knowing that each $M(s)$ omits each Δ_m or is atomic is enough to conclude that M omits each Δ_m or is atomic.

6 Applications I - When does an atomic model of size $\aleph_1$ imply one of $\beth_1$?

In this section, we use the generalized Henkin method to find a number of sufficient conditions on T for which the existence of an atomic model of size $\aleph_1$ implies the existence of an atomic model of size $\beth_1$. In the first subsection, we show that if every set is definably closed, a very straightforward argument leads from a *countable*[9] model to one in the continuum. In particular, there is no need for the Y-variables from our general formulation. In the second and third subsections we formalize the conditions used in the first in terms of combinatorial geometry and get a general result which specializes to the goal which motivated this project: In pseudo-minimal theories [BLS16], the existence of an uncountable, atomic model implies one of size continuum. Then, in the fourth subsection, we move to material that requires much more background and show how the arguments of Hrushovski and Shelah in [HS91] can be put into our framework. There, they prove that if a countable, superstable theory T has an atomic model of size $\aleph_1$, then it has an atomic model of size $\beth_1$.

6.1 Theories with trivial dcl

In a series of papers, e.g., [AFP16], Ackerman, Freer, and Patel found that classes of models of theories with trivial definable closure have some very desirable properties. Here we note that such theories behave exceptionally well with respect to the Henkin constructions described in this paper. In particular, we will see that the **Henkin** and **Splitting** conditions will be easily satisfied in any model of such a theory.

We begin with a pair of classical definitions.

Definition 6.1.1. Given an L-structure M and subset $A \subseteq M$, an element $b \in M$ is *A-definable* if there is a formula $\varphi(x, \bar{a})$ with $\bar{a}$ from A for which b is the only solution in M. The *definable closure of A*, $\mathrm{dcl}(A)$ is the set of A-definable elements of M.

Similarly, $b \in M$ is *A-algebraic* if there is an integer k and a formula $\varphi(x, \bar{a})$ such that $M \models \varphi(b, \bar{a})$ and $M \models \exists^{=k} x \varphi(x, \bar{a})$. The *algebraic closure of A*, $\mathrm{acl}(A)$, is the set of A-algebraic elements of M.

[9] Using Theorem 6.3.2, it is easy to see any structure with trivial definable closure is $L_{\omega_1,\omega}$-equivalent to an uncountable structure.

Clearly, $A \subseteq \mathrm{dcl}(A) \subseteq \mathrm{acl}(A)$ for any subset $A \subseteq M$. We distinguish structures for which both of these closures are trivial.

Definition 6.1.2. Fix a countable language L. An L-structure M has *trivial definable closure* (is dcl-trivial) if $\mathrm{dcl}(A) = A$ for every subset $A \subseteq M$.

Note that this is very different notion from the usual usage of a trivial closure relation in combinatorial geometry. Note also that dcl-triviality is distinct from atomicity. In particular, the theory of countably many independent unary relations is dcl-trivial but has no atomic models.

It is clear that any dcl-trivial structure is infinite, and that dcl-triviality is a property of the theory of M, i.e., if N is elementarily equivalent to M, then N is dcl-trivial if and only if M is.

The *key property* of a dcl-trivial structure M is easy to see: if $M \models \exists u\varphi(u, \bar{c}) \wedge u \notin \bar{c}$, then $\varphi(u, \bar{c})$ has infinitely many solutions in M. From the key property it is easily seen that dcl-triviality of M is equivalent to $\mathrm{acl}(A) = A$ for every $A \subseteq M$. In what follows, we will see that dcl-triviality has many equivalent formulations. A roster of equivalents is given in Fact 7.2.9.

Constructing models of theories with trivial dcl is by far the most straightforward example of our technique, which justifies our considering it first. The simplicity comes from the fact that we do not require any Y-variables! But, we must doubly index the x's as $x_{\eta,i}$.

Definition 6.1.3. Let N be any L-structure. Suppose $\psi(\overline{x}, \bar{y})$ is an L-formula with $\mathrm{lg}(\overline{x}) = k$. For any $\bar{b}$ from N, call the definable subset $\psi(N^k, \bar{b})$ of N^k *non-degenerate* if there exists some $\bar{a} \in \psi(N^k, \bar{b})$ with $\{a_1, \ldots, a_k\}$ pairwise distinct and disjoint from $\bar{b}$.

Theorem 6.1.4. *Suppose M is a* dcl*-trivial structure in a countable language L. There is a model N elementarily equivalent to M of size continuum that satisfies:*

1. *The universe of N is indexed as $2^\omega \times \omega$;*

2. *The universe of N can be partitioned as $N = \bigcup_{i\in\omega} A_i$, where, for each i, $A_i = \{a_{\eta,i} : \eta \in 2^\omega\}$ is an asymptotically similar subset;*[10]

[10] In fact, for every finite, strictly increasing sequence $t = (i_1, i_2, \ldots, i_k)$ from ω, the sequences $\{\bar{a}_{\eta,t} : \eta \in 2^\omega\}$ (where $\bar{a}_{\eta,t} = (a_{\eta,i_1}, \ldots, a_{\eta_{i_k}})$) is an asymptotically similar set of k-tuples.

3. *With respect to the natural Polish topology*[11] *on* $2^\omega \times \omega$*, for every* k*, every definable subset of* N^k *is a finite boolean combination of open sets of* $(2^\omega \times \omega)^k$*, with the product topology.*

4. *If we place the usual measure*[12] *on* $2^\omega \times \omega$*, then for every* k*, every non-degenerate definable subset of* N^k *has positive measure (with respect to the product measure on* $(2^\omega \times \omega)^k$*.*

5. *If, in addition,* M *is atomic, then we can insist that* N *be atomic as well;*

6. *More generally, if* $\{\Delta_m : m \in \omega\}$ *is a countable set of types omitted by* M*, then we can insist that* N *omits each* Δ_m *as well.*

Remark 6.1.5. In fact, in (3) we can say more – the bound on the size of the boolean combination depends only on k, and not on either the language L or the choice of L-structure. That is, there is a function $k \mapsto n(k)$ with the property that for every countable L and every dcl-trivial L-structure M, the associated N has the property that every definable subset of N^k is a boolean combination of at most $n(k)$ open subsets.

Proof of Theorem 6.1.4: Fix a dcl-trivial M. We take $Z = X$, where X is doubly indexed as $\{x_{\eta,i} : \eta \in 2^\omega, i \in \omega\}$. To define our set of commitments, first let $\mathcal{D}_0$ consist of all L-formulas $\varphi(\overline{w})$ that imply $w_j \neq w_{j'}$ for distinct $j \neq j'$ that are consistent with $T = Th(M)$. For each fmac A of $2^{<\omega}$, let $Z_A = \{x_{a,i} : a \in A, i \in \omega\}$. Then, for each such A, let the set of A-commitments $\mathbb{P}_A$ consist of all Z_A-instantiations of formulas $\varphi(\overline{w}) \in \mathcal{D}_0$ by a tuple $\bar{z}$ of *distinct* elements of Z_A.

Let $(\mathbb{P}, \leq)$ be the poset with universe $\mathbb{P} = \bigcup\{\mathbb{P}_A\colon\ A \text{ an fmac of } 2^{<\omega}\}$ and where $\leq$ is the extension relation from Section 5. We show that **Completeness**, **Henkin witnesses**, and **Splitting** conditions follow easily: Fix any fmac A and any A-commitment[13] $\varphi(\overline{x}) \in \mathbb{P}_A$. As $\varphi(\overline{x})$ is consistent with $Th(M)$, choose $\bar{c}$ from M such that $M \models \varphi(\bar{c})$.

[11] The basis consists of sets of the form $U_a \times \{i\}$where U_a are as in Section 4.

[12] For any basic open $U_a \subseteq 2^\omega$ with $|a| = n$ let $\mu(U_a) = \frac{1}{2^n}$ and then extend to $2^\omega \times \omega$ by letting $\mu(U_a \times \{i\}) = \frac{1}{2^{n+i+1}}$. In fact, if we regard the base set as the locally compact group given by pointwise addition on ω copies of Z_2^ω, this is a Haar measure.

[13] We sometimes abuse notation by identifying $\mathbb{P}_A$ with the formulas that occur as second coordinates of the pairs.

Completeness: Given a $\psi(\bar{z})$, where $\bar{z}$ is a subsequence of $\overline{x}$, we will show it is decided. Let $\bar{b}$ be the corresponding subsequence of $\bar{c}$. Now, if $M \models \psi(\bar{b})$, then put $\varphi^* := \varphi(\overline{x}) \wedge \psi(\bar{z})$, and put $\varphi^* := \varphi(\overline{x}) \wedge \neg\psi(\bar{z})$ otherwise.

Henkin witnesses: We must satisfy the condition for an arbitrary $\theta(w, \bar{z})$ with $\bar{z}$ a subsequence of $\overline{x}$. Let $t := t(\bar{z})$ be the set of $a \in A$ such that for some i, a variable $x_{a,i}$ appears in $\bar{z}$. As above, let $\bar{b}$ be the subsequence of $\bar{c}$ associated to $\bar{z}$. There are three cases. First, if $M \models \neg\exists w\theta(w, \bar{c})$, then, put $\varphi^* := \varphi(\overline{x}\overline{y}) \wedge \neg\exists w\theta(w, \bar{z})$. Then $\bar{c}$ witnesses that φ^* is an A-commitment and it is evident that (A, φ^*) extends (A, φ).

Second, suppose $M \models \theta(c, \bar{c})$ for some $c \in \bar{c}$. Let z^* be the (unique) element of $\bar{z}$ corresponding to c. Then $\varphi^* := \varphi(\overline{x}) \wedge \theta(z^*\bar{z})$ is in $\mathbb{P}_A$ and extends $\varphi(\overline{x})$.

Finally, suppose $M \models \exists u\theta(u, \bar{c}) \wedge \bigwedge u \notin \bar{c}$. Then, by the key property of dcl-triviality, choose $b^* \in M \setminus \bar{c}$ such that $M \models \theta(b^*, \bar{c})$. Choose any $a \in t(\bar{z})$ and $j \in \omega$ such that $x_{a,j} \notin \overline{x}$ and put

$$\varphi(x_{a,j}\overline{x}) := \varphi(\overline{x}) \wedge \psi(x_{a,j}\bar{z}) \wedge \bigwedge x_{a,j} \notin \overline{x}$$

Then $b^*\bar{c}$ witnesses that $\varphi^* \in \mathbb{P}_A$, which visibly extends φ.

Splitting: Choose any $a \in A$. To handle this case, we start with a Claim, whose proof is an easy induction on k; the key property yields the case $k = 1$:

Claim. For every $k \geq 1$, for every $\varphi(\overline{x}) \in \mathcal{D}_0$, and for every partitioning of $\overline{x} = \bar{u}\bar{v}$ with $\lg(\bar{u}) = k$, then for every $\bar{b}$ from M such that $M \models \exists\bar{u}\varphi(\bar{u}, \bar{b})$, there is an infinite, pairwise disjoint set $\{\bar{c}_j : j \in \omega\} \subseteq M^k$ of realizations of $\varphi(\bar{u}, \bar{b})$.

Given the Claim, partition the variables of $\varphi(\overline{x})$ into two disjoint subsequences $\overline{x} = \overline{x}_a\overline{x}^*$, where $\overline{x}_a$ consists of all $x_{a,i} \in \overline{x}$, while $\overline{x}^*$ consists of all $x_{a',i} \in \overline{x}$ with $a' \neq a$. This partition induces a partition of our realizing sequence $\bar{c}$ into $\bar{c}_a\bar{b}$, where $\bar{c}_a$ corresponds to $\overline{x}_a$, while $\bar{b}$ corresponds to $\overline{x}^*$. Put

$$\varphi^*(\overline{x}_{a\hat{\ }0}, \overline{x}_{a\hat{\ }1}, \overline{x}^*) := \varphi(\overline{x}_{a\hat{\ }0}, \overline{x}^*) \wedge \varphi(\overline{x}_{a\hat{\ }1}, \overline{x}^*) \wedge \text{ `}\overline{x}_{a\hat{\ }0}, \overline{x}_{a\hat{\ }1}, \overline{x}^* \text{ are distinct'}$$

Then the Claim implies that $(A^{*a}, \varphi^*) \in \mathbb{P}_{A^{*a}}$, and is as required.

Now, with our density conditions satisfied, the existence of a model N follows from Theorem 5.4. By our choice of $\mathcal{D}_0$, the congruence $\sim$ on $Z = X$

is trivial, which establishes Clause 1) and the partition of Clause 2). The remaining Clauses are established by the properties guaranteed by Theorem 5.4 and the footnotes.

6.2 Sufficient pregeometries

In this and the following subsection we study the effect of having an atomic model that is equipped with a well behaved closure relation. In this subsection we give a sufficient set of conditions on a closure relation of an atomic model (M, cl) to allow for the construction of an elementarily equivalent atomic model of size continuum. As an application, in the next subsection we prove a new result: among pseudo-minimal theories, the existence of an uncountable, atomic model implies one of size continuum.

Although we have cast our results in terms of the existence of atomic models, they translate to complete sentence of $L_{\omega_1,\omega}$ as in Section 3.3 (equivalently for countable, first order theories that omit a given type).

Definition 6.2.1. Let M be any L-structure. A *formula-based closure relation on M* is a function $\mathrm{cl} : \mathcal{P}(M) \to \mathcal{P}(M)$ satisfying for all $A, B \subseteq M$, $A \subseteq \mathrm{cl}(A)$; $A \subset B$ implies $\mathrm{cl}(A) \subseteq \mathrm{cl}(B)$; $\mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$; and whenever $a \in \mathrm{cl}(B)$, then there is a finite tuple $\bar{b}$ from B and a formula $\varphi(x, \bar{y}) \in \mathrm{tp}(a\bar{b})$ such that $a' \in \mathrm{cl}(\bar{b}')$ whenever $M \models \varphi(a', \bar{b}')$.

Formula-based closure relations abound in model theory. Examples include *equality* $(M, =)$, where $\mathrm{cl}(A) = A$ for all $A \subseteq M$, *definable closure* (M, dcl), and *algebraic closure* (M, acl). Additionally, in the next subsection we introduce *pseudo-algebraic closure* (M, pcl), which is well behaved whenever M is atomic. In order to apply our methods, we need our formula-based closure relation to satisfy more properties.

Definition 6.2.2. Consider a formula-based closure relation (M, cl) on an arbitrary infinite L-structure. We call (M, cl) *sufficient* if the following additional conditions hold:

1. 'Exchange:' i.e., if $a \in \mathrm{cl}(Bc) \setminus \mathrm{cl}(B)$, then $c \in \mathrm{cl}(Ba)$;

2. 'Extendible[14]:' There is $a \in M \setminus \mathrm{cl}(\emptyset)$; and

[14] If any of dcl, acl, or pcl are not extendible, the Scott sentence of M has exactly one model.

3. 'Weak homogeneity:' For all finite $\bar{b}$ and L-formulas $\varphi(w, \bar{b})$, if there is $a \notin \mathrm{cl}(\bar{b})$ with $M \models \varphi(a, \bar{b})$, then for every finite $E \subseteq M$, there is $a' \notin \mathrm{cl}(E)$ that also satisfies $M \models \varphi(a', \bar{b})$.

A closure relation that satisfies Exchange is also known as a *pregeometry* or a *matroid*. It is well known that pregeometries give rise to a well behaved notion of dimension. In particular, for any set B, any two maximal independent subsets of $\mathrm{cl}(B)$ have the same cardinality. One of many introductions to the role of

Remark 6.2.3. We say $\bar{a}$ is *independent over* E if for every $i < \lg(\bar{a})$, $a_i \notin \mathrm{cl}(\bar{a} - \{a_i\} \cup E)$. A routine induction shows that the 'Weak homogeneity' condition implies that for every n, every $\psi(\overline{w}, \bar{b})$, if there is an n-tuple $\bar{a}$ independent over $\bar{b}$ with $M \models \psi(\bar{a}, \bar{b})$, then for every finite E, there is $\bar{a}'$ independent over E with $M \models \psi(\bar{a}', \bar{b})$. Also, coupled with 'Extendible', we conclude that M contains an infinite independent subset I. Moreover, for any L-formula $\varphi(w, \bar{b})$, either $\varphi(M, \bar{b}) \subseteq \mathrm{cl}(\bar{b})$, or for every finite set E, $\varphi(M, \bar{b})$ contains an infinite, E-independent subset.

Examples of sufficient pregeometries are common. A structure $(M, =)$ has a sufficient pregeometry if and only if M has trivial dcl. If T is strongly minimal, weakly minimal, o-minimal, or has SU-rank 1, then (M, acl) is a pregeometry for any model of T. Moreover, an easy compactness argument shows that any (infinite) model M of such a theory has a proper, elementary extension N for which (N, acl) is sufficient. In the next subsection we prove that whenever a pseudo-minimal theory has an uncountable atomic model, then (M, pcl) is sufficient for every atomic model. For now, we content ourselves with the following result.

Theorem 6.2.4. *Suppose* (M, cl) *is a sufficient pregeometry. Then there is a Borel model* $N \equiv M$ *of size continuum with a* cl*-independent, asymptotically similar subset* $\{a_\eta : \eta \in {}^\omega 2\}$ *from* N*. Moreover, if* M *is atomic (with respect to* $Th(M)$*) then we may additionally choose* N *to be atomic. More generally, if* $\{\Delta_m(\overline{w}_m) : m \in \omega\}$ *is a countable set of partial types, each of which is omitted in* M*, then we may additionally require that* N *omits every* Δ_m*.*

Proof. In this application, it is helpful to doubly index the X-variables. That is, take as variables $X = \{x_{\eta,i} : \eta \in 2^\omega, i \in \omega\}$, as usual, $Y = \{y_{s,i} : s \subseteq 2^\omega \text{ finite}, i \in \omega\}$ and $Z = X \cup Y$. The double indexing of the X-variables is needed since a typical model (e.g., some M_η) may have an infinite, independent subset.

As **notation** , for any fmac A, any non-empty subset $t \subseteq A$, and any $\overline{x} \in X_A$, $\overline{x}_t$ denotes the subsequence of $\overline{x}$ from X_t, i.e., an element $x_{a,i} \in \overline{x}$ is an element of $\overline{x}_t$ if and only if $a \in t$. Similarly, for any $\bar{y} \in Y_A$, $\bar{y}_t$ is the subsequence of $\bar{y}$ from Y_t, i.e., for $y_{s,i} \in \bar{y}$, $y_{s,i} \in \bar{y}_t$ if and only if $s \subseteq t$.

For any fmac A, let $\mathbb{P}_A$ denote all Z_A-instantiated formulas $\varphi(\overline{x}, \overline{y})$ where $\overline{x} \in X_A$, $\bar{y} \in Y_A$ and there are sequences $\bar{c}, \bar{b}$ from M satisfying:

1. $M \models \varphi(\bar{c}, \bar{b})$;
2. $\bar{c}$ is cl-independent; and
3. For each $t \subseteq A$, $M \models \forall \overline{x} \forall \overline{y} (\varphi(\overline{x}, \overline{y}) \rightarrow \bar{y}_t \subseteq \mathrm{cl}(\overline{x}_t))$ (cf., 'Formula-basedness')

As usual, let $(\mathbb{P}, \leq)$ be the poset with universe

$$\mathbb{P} = \{(A, \varphi) : A \text{ is a fmac and } \varphi \in \mathbb{P}_A\}$$

and $\leq$ is the usual extension relation. We argue that $(\mathbb{P}, \leq)$ satisfies **Completeness**, **Henkin witnesses**, and **Splitting.**

Fix an fmac A and an A-commitment $(A, \varphi(\overline{x}, \overline{y})) \in \mathbb{P}_A$. Choose finite tuples $\bar{c}, \bar{b}$ from M witnessing that $\varphi \in \mathbb{P}_A$.

Completeness: Choose any $\psi(\bar{z})$ with $\bar{z}$ from Z_A, which we may assume is a subsequence of $\overline{x}\overline{y}$. Let $\bar{d}$ be the corresponding subsequence of $\bar{c}\bar{b}$. There are now two cases: If $M \models \psi(\bar{d})$, then put $\varphi^*(\overline{x}\overline{y}) := \varphi(\overline{x}\overline{y}) \wedge \psi(\bar{z})$; and put $\varphi^*(\overline{x}\overline{y}) := \varphi(\overline{x}\overline{y}) \wedge \neg\psi(\bar{z})$ otherwise. In either case, the same pair $\bar{a}\bar{b}$ demonstrate that $\varphi^* \in \mathbb{P}_A$.

Henkin witnesses: Choose $\theta(w, \bar{z})$ with $\bar{z}$ from Z_A, which we may again assume is a subsequence of $\overline{x}\overline{y}$. As above, let $\bar{d}$ be the subsequence of $\bar{c}\bar{b}$ corresponding to $\bar{z}$, and in the notation of Definition 5.2 as amplified just above, let $t = t(\bar{z}) \subseteq A$. There are now three cases. First, if $M \models \neg\exists w\theta(w, \bar{d})$, then put $\varphi^*(\overline{x}\overline{y}) := \varphi(\overline{x}\overline{y}) \wedge \neg\exists w\theta(w, \bar{z})$.

Second, suppose there is $h \in \mathrm{cl}(\bar{c}_t)$ such that $M \models \theta(h, \bar{d})$. By 'formula-basedness' choose a formula $\delta(w, \overline{x}_t) \in \mathrm{tp}(h, \bar{c}_t)$ such that any realization of $\delta(w, \bar{c}_t)$ in M implies $w \in \mathrm{cl}(\bar{c}_t)$. Choose i such that $y_{t,i} \notin \bar{y}$. Put

$$\varphi^*(\overline{x}, \bar{y}y_{t,i}) := \varphi(\overline{x}, \overline{y}) \wedge \theta(y_{t,i}, \bar{z}) \wedge \delta(y_{t,i}, \overline{x}_t)$$

That $\varphi^* \in \mathbb{P}_A$ is witnessed by appending h to $\bar{b}_t$.

Third, suppose there is $h \in M \setminus \mathrm{cl}(\bar{c}_t)$ such that $M \models \theta(h, \bar{d})$. Then, clearly, $\{h\} \cup \bar{c}_t$ is independent. Choose any $i \in \omega$ such that $x_{t,i} \notin \overline{x}_t$. Put

$$\varphi^*(x_{t,i}\overline{x}, \overline{y}) := \varphi(\overline{x}, \overline{y}) \wedge \theta(x_{t,i}, \bar{z})$$

By Weak Homogeneity choose $c^* \notin \mathrm{cl}(\bar{c}\bar{b})$ with $M \models \theta(c^*, \bar{d})$. As φ^* is witnessed by $c^*\bar{a}\bar{b}$, it follows that $\varphi^* \in \mathbb{P}_A$ and extends φ.

Splitting: Choose any $a \in A$ and let $A^- = A \setminus \{a\}$. Partition the variables of $\bar{z} = \overline{x}\overline{y}$ into four disjoint subsequences:

- $\overline{x}_a$ is the subsequence of $\overline{x}$ consisting of all $x_{a,i} \in \overline{x}$;
- $\overline{x}_0$ is the subsequence of $\overline{x}$ consisting of all $x \in X_{A^-}$;
- $\bar{y}_a$ is the subsequence of $\bar{y}$ consisting of all $y_{s,i} \in \bar{y}$ for which $a \in s$; and
- $\bar{y}_0$ is the subsequence of $\bar{y}$ consisting of all $z \in Z_{A^-}$ (i.e., whose coordinates do not mention a).

As notation, let $\bar{c}_a, \bar{c}_0, \bar{b}_a, \bar{b}_0$ denote the subsequences of $\bar{c}\bar{b}$ corresponding to $\overline{x}_a, \overline{x}_0, \overline{y}_a, \overline{y}_0$, respectively. Put $\psi(\overline{x}_a, \overline{x}_0, \overline{y}_0) := \exists \overline{y}_a \varphi$. Then $M \models \psi(\bar{c}_a, \bar{b}_0, \bar{c}_0)$ as witnessed by $\bar{b}_a$. Furthermore, $\bar{c}_a\bar{c}_0$ form a partition of $\bar{c}$ and hence are independent. Thus, by condition 3) $\bar{b}_a\bar{b}_0 \subseteq \mathrm{cl}(\bar{c}_a\bar{c}_0)$ and $\bar{c}_a$ is independent over $\mathrm{cl}(\bar{b}_0\bar{c}_0)$.

By Remark 6.2.3 choose $\bar{c}'_a$ from M realizing $\psi(\overline{x}_a, \bar{b}_0\bar{c}_0)$ and independent from all of $\bar{c}\bar{b}$. In particular, $\bar{c}'_a$ is disjoint from $\bar{c}_a$. By choice of ψ, choose $\bar{b}'_a$ from M such that $M \models \varphi(\bar{c}'_a, \bar{c}_0, \bar{b}'_a, \bar{b}_0)$. It follows that $\bar{b}'_a \subseteq \mathrm{cl}(\bar{c}'_a\bar{c}_0\bar{b}_0)$. It is easily checked that these tuples witness: $[h_0(\varphi) \wedge h_1(\varphi) \wedge \overline{x}_{a\hat{}0} \cap \overline{x}_{a\hat{}1} = \emptyset] \in \mathbb{P}_{A^{*a}}$.

6.3 Pseudominimal Theories

In a series of papers, the authors and Shelah have attempted to determine whether every $\aleph_1$-categorical, complete sentence Φ of $L_{\omega_1,\omega}$ has a model of size continuum. By the reductions in Subsection 3.3, this is equivalent to asking whether a complete first order theory T that has a unique atomic model of size $\aleph_1$ must also have an atomic model of size continuum.

To analyze this problem, in [BLS16], we introduced a new notion of closure, which we dubbed *pseudo-closure*, shortening pseudo-algebraic closure, that is appropriate for the study of atomic models of a first order theory. We proved that if pseudo-closure fails exchange in a strong way on the class of atomic models of a theory T then T has $2^{\aleph_1}$ atomic models of cardinality $\aleph_1$. We give a slightly simplified account of pseudo-minimality which is adequate for the applications. Here we show that if T has an uncountable atomic model that is pseudo-minimal, then there is an atomic model of T in the continuum.

Definition 6.3.1. Let M be an atomic model and suppose $a, \bar{b}$ are from M. We say a is *pseudo-algebraic over* $\bar{b}$ *in* M, written $a \in \mathrm{pcl}(\bar{b})$, if every elementary substructure $N \preceq M$ that contains $\bar{b}$ also contains a.

We showed in [BLS16] that pseudo-algebraicity in atomic models is formula-based and a property of the theory as opposed to a particular model. That is, if M and M' are elementarily equivalent atomic models, $\bar{a}, \bar{b}$ and $\bar{a}', \bar{b}'$ are from M and M', respectively, whose pairs realize the same complete formula, then $\bar{a} \in \mathrm{pcl}(\bar{b})$ in M if and only if $\bar{a}' \in \mathrm{pcl}(\bar{b})$ in M'. Also, Lemma 2.6 of [BLS16] implies that if M is atomic, then (M, pcl) satisfies the 'Weak homogeneity' clause from Definition 6.2.2.

Using this notion we can immediately add a clause to an old theorem of Vaught.

Lemma 6.3.2. *Let T be a complete theory in a countable language that has an atomic model. The following notions are equivalent:*

- *T has an uncountable atomic model;*
- *the countable atomic model has a proper atomic extension;*
- *the countable atomic model is not minimal; and the new*
- $\mathrm{pcl}(\emptyset) \neq M$ *for some/every atomic model.*

Definition 6.3.3. Let M be an atomic model and suppose T satisfies the conditions of Lemma 6.3.2. We say that T is *pseudominimal* if (M, pcl) satisfies Exchange for some/every atomic model M of T. That is, for every finite set C from M and elements $a, b \in M$, if $b \in \mathrm{pcl}(Ca)$ but $b \notin \mathrm{pcl}(C)$, then $a \in \mathrm{pcl}(Cb)$.

Thus, a complete theory T satisfying the hypotheses of Lemma 6.3.2 is pseudominimal if and only if (M, pcl) is a sufficient pregeometry for some/every atomic model M of T.

The following new Theorem is a culmination of our previous results. It follows immediately from Lemma 6.3.2, the note above, and Theorem 6.2.4.

Theorem 6.3.4. *If a countable first order theory T has an atomic pseudominimal model M of cardinality $\aleph_1$ then there is an atomic pseudominimal model N of T with cardinality $2^{\aleph_0}$.*

Equivalently, if the models of a complete sentence Φ in $L_{\omega_1,\omega}$ are pseudominimal and Φ has an uncountable model, it has a model in the continuum.

Whereas Theorem 6.3.4 is of general interest, we note a special case. It is an easy exercise to prove that any weakly minimal theory T with an uncountable atomic model is pseudominimal. Thus, Theorem 6.3.4 gives a proof that such a theory has an atomic model of size continuum (a second is Theorem 6.4.1).

As an example of pseudominimality, Zilber [Zil05, Bal09] introduced the abstract notion of a quasiminimal (excellent) class and proved such classes are categorical in all uncountable powers. In general, these classes are axiomatized in $L_{\omega_1,\omega}(Q)$ ([Kir10]) and the quasiminimal closure is distinct from our notion of pcl. However, in some cases, most notably [BZ11], the study of covers of certain algebraic groups e.g. [BZ11, Bay09], the countability of the quasiminal closure is expressible in $L_{\omega_1,\omega}$ and then pcl = qcl.

6.4 Stable and superstable theories

Stable theories give rise to a well-behaved notion of independence, namely non-forking. Using this tool in conjunction with the methods of this paper, Hrushovski and Shelah [HS91] obtain the following transfer theorem:

Theorem 6.4.1. *Suppose N is an uncountable model of a superstable theory T in a countable language. Then there is an atomic model M of T of size continuum that has an asymptotically similar subset $\{a_\eta : \eta \in 2^\omega\}$.*

We sketch their proof of Theorem 6.4.1 using the technology described here. In fact, in [HS91] they prove more – If $\{\Delta_m(\overline{w}_m) : m \in \omega\}$ is any countable set of partial types and there is an uncountable model N of a countable, superstable theory T omitting each Δ_m, then there is a model M of size continuum, again with an asymptotically similar subset, that also omits each Δ_m. As well, using the same machinery they obtain the same conclusion for a countable stable theory, at the cost of requiring the original model N to have size $\aleph_{\omega+1}$.

By employing the extensive calculus of non-forking, Shelah has gleaned many structural consequences from his notion of a *stable system* of models.

Definition 6.4.2. Let I be any non-empty index set. A *stable system* of countable models of T is a set $\{M(s) : s \in [I]^{<\omega}\}$ of countable models of T satisfying:

- If $s \subseteq t$, then $M(s) \preceq M(t)$;
- For all $s, t \in [I]^{<\omega}$, then $M(s)$ and $M(t)$ are independent (i.e., do not fork) over $M(s \cap t)$.

A primary tool for construction stable systems of models is *domination.* That is, given a pair of models $M \preceq M'$ and a subset $B \subseteq M'$, we say B *dominates* M' *over* M if, for any set X (in some larger model), if X is independent from B over M, then X is independent from M' over M. As we are working over models in a stable theory, a sufficient condition for domination is Lachlan's notion [Lac72] of locally atomic models, ℓ-atomicity:

Definition 6.4.3. Given a set B, a complete type $p \in S_n(B)$ is *locally (ℓ-isolated)* if, for every partitioned formula $\varphi(\overline{x}, \overline{y})$, there is a formula $\psi(\overline{x}) \in p$ such that $\psi(\overline{x}) \vdash \varphi(\overline{x}, \bar{b})$ for every $\varphi(\overline{x}, \bar{b}) \in p$. We call a model M' *ℓ-atomic over B* if, for every finite $\bar{a}$ from M', $\mathrm{tp}(\bar{a}/B)$ is ℓ-isolated.

A fundamental fact is that for stable theories, if $M \subseteq B$ and if M' is ℓ-atomic over B, then M' is dominated by B over M.

Hrushovski and Shelah's proof of Theorem 6.4.1 breaks into two pieces. The first part, which uses some highly technical stability-theoretic machinery (including the existence of definable groups in some instances) states that one can find a 'very rich' stable system indexed by $I = \omega_1$ of elementary substructures of any uncountable model N of a superstable theory T.

Theorem 6.4.4. *[HS91] Let N be an uncountable model of a countable, superstable theory T. There is a stable system $\{M(s) : s \in [\omega_1]^{<\omega}\}$ of countable, elementary substructures of N and an independent subset $C = \{c_i : i \in \omega_1\}$ over $M(\emptyset)$ of N that satisfy:*

1. *For each $i \in \omega_1$, $c_i \in M(\{i\})$ and $M(\{i\})$ is ℓ-atomic over Mc_i;*
2. *For each $i \in \omega_1$ and $\theta(x, \bar{b}) \in \mathrm{tp}(c_i/M(\emptyset))$, there are infinitely many $j \in \omega_1$ such that $M(\{j\}) \models \theta(c_j, \bar{b})$; and*

3. For $|s| \geq 2$, $M(s)$ is ℓ-atomic over $\bigcup\{M(t): t \subsetneq s\}$.

As this theorem is rather technical, we only sketch the argument here and use some unexplained notation.

Proof sketch. Without loss, we may assume N has cardinality $\aleph_1$. Fix an enumeration $\langle a_i : i \in \omega_1 \rangle$ of N. For each $i \in \omega_1$, let $A_i = \{a_j : j < i\}$ and let $p_i = tp(a_i/A_i)$. As each p_i is based on a finite set, for each i there is some $j < i$ such that p_i is based on A_j. By Fodor's Lemma, there is some j^* and a stationary subset $S \subseteq \omega_1$ such that for each $i \in S$, $i > j^*$ and p_i is based on A_{j^*}. Fix such a j^* and put $B := A_{j^*}$. So B is countable, and by reindexing S, we have an uncountable set $C = \{c_i : i \in \omega_1\}$ that is independent over B.

Next, choose a countable $M \preceq N$ such that $B \subseteq M$ and M is an *na*-substructure of N. Using superstability, by removing at most countably many of the c_i's we obtain that the remaining, uncountably many elements are independent over M.

Now that we have chosen M and I, it remains to construct our stable system $\langle M(s) : s \in [\omega_1]^{<\omega} \rangle$. But this follows immediately by successive applications of the Corollary on page 302 of [HS91].

The second part of the proof of Theorem 6.4.1 can be proved using the technology of this paper. For this half, only stability is needed.

Theorem 6.4.5. *Suppose T is a countable, stable theory and $\{M(s) : s \in [\omega_1]^{<\omega}\}$ is a stable system of countable elementary submodels of an atomic model N satisfying Clauses (1)-(3) of Theorem 6.4.4. Then there is a Borel, atomic model N_1 of size continuum with an asymptotically similar subset $\{a_\eta : \eta \in 2^\omega\}$. More generally, if N omits a countable set $\{\Delta_m : m \in \omega\}$ of types, then N_1 can be chosen to omit each Δ_m.*

Proof. For this application, we take our set Z of variable symbols to be $X \cup Y$, where $X = \{x_\eta : \eta \in 2^\omega\}$ and $Y = \{y_{s,i} : s \in [2^\omega]^{<\omega}, i \in \omega\}$.

Choose any fmac $A \subseteq 2^{<\omega}$ with an enumeration $\langle a_j : j \in A \rangle$. Suppose that $f : A \to \omega_1$ is any injective mapping. Any such f describes a finite tuple $\bar{c}_f := \langle c_{f(j)} : j \in A \rangle$ from the distinguished independent set $C = \{c_i : i \in \omega_1\}$. Also, f extends to a map $f : \mathcal{P}(A) \to [\omega_1]^{<\aleph_0}$ by $f(t) := \{c_{f(j)} : j \in t\}$.

With this notation, define the set $\mathbb{P}_A$ of A-commitments to be the set of instantiated Z_A-formulas $\varphi(\overline{x}, \bar{z})$, where $\bar{z} := \langle \bar{y}_s : s \subseteq A \rangle$ and each tuple $\bar{y}_s$ is from $\{y_{s,i} : i \in \omega\}$, for which there is some injective $f : A \to \omega_1$ and tuples $\langle \bar{b}_s : s \subseteq A \rangle$ from $M(f(s))$ so that $N \models \varphi(\bar{c}_f, \bar{b}_s : s \subseteq A)$. As usual, let $(\mathbb{P}, \leq)$ be the partial order where $\mathbb{P} = \bigcup\{\mathbb{P}_A : A \text{ an fmac}\}$ and $\leq$ is defined as in Section 4. As the given model N and hence each of the submodels $M(s)$ omit each Δ_m, the **Omitting** Δ_m conditions are easily verified. As well, the verifications of the density conditions **Completeness** and **Henkin witnesses** are straightforward. For both, fix an fmac A and an A-commitment $\varphi(\overline{x}, \overline{y}) \in \mathbb{P}_A$. Choose an injective function $f : A \to \omega_1$ and tuples $\bar{b}_s$ from $M(f(s))$ such that $M(f(A)) \models \varphi(\bar{c}_f, \bar{b}_s : s \subseteq A)$.

Completeness: Choose any instantiated Z_A-formula $\psi(\bar{z})$ and partition its variables as $\psi(\overline{x}, \bar{y}_s : s \subseteq A)$. By adding dummy variables to both φ and ψ, we may assume they have the same instantiated variables. To decide how to extend φ, we simply appeal to $M(f(A))$. On one hand, if $M(f(A)) \models \psi(\bar{c}_f, \bar{b}_s : s \subseteq A)$, then put $\varphi^* := \varphi \wedge \psi$; put $\varphi^* := \varphi \wedge \neg\psi$ otherwise.

Henkin witnesses: Choose any instantiated Z_A-formula $\theta(w, \bar{z})$ with w free. In the notation of Definition 5.2, put $t := t(\bar{z})$. Then the subsequence $\bar{d}$ of $\langle \bar{c}_f, \bar{b}_s : s \subseteq A\rangle$ corresponding to $\bar{z}$ is contained in $M(f(t))$. As above, there are two cases. If $M(f(A)) \models \neg\exists w\theta(w, \bar{d})$, then put $\varphi^* := \varphi \wedge \neg\exists w\theta(w, \bar{z})$. Otherwise, append a new element $y_{t,j}$ to $\bar{y}_t$, forming $\bar{y}'_t$, and put $\varphi^* := \varphi \wedge \theta(y_{t,j}, \bar{z})$. As $M(f(t)) \preceq M(f(A))$, there is $b^* \in M(f(t))$ witnessing $\theta(w, \bar{d})$. This extra element witnesses that $\varphi^* \in \mathbb{P}_A$.

By contrast, the verification of **Splitting** is more involved, and requires new ideas. As above, fix an enumerated fmac $A = \langle a_i : i < n\rangle$ and an injective $f : A \to \omega_1$ that witnesses that $\varphi(\overline{x}, \bar{z}) \in \mathbb{P}_A$. Choose an arbitrary $a \in A$, but to ease notation, suppose that $a = a_0$ and choose $\varphi(\overline{x}, \bar{z}) \in \mathbb{P}_A$. As notation, let $A^- = A \setminus \{a_0\}$, let $A_0 = A^- \cup \{a\hat{\ }0\}$ and $A_1 = A^- \cup \{a\hat{\ }1\}$. Thus, $A^{*a} = A_0 \cup A_1$ and the liftings $h_0, h_1 : A \to A^{*a}$ map onto A_0, A_1, respectively. Fix an enumeration $\langle s_i : i < 2^n\rangle$ of $\mathcal{P}(A)$ that satisfies (I) $i \leq j$ whenever $s_i \subseteq s_j$ and (II) the initial segment $\langle s_i : i < 2^{n-1}\rangle$ enumerates $\mathcal{P}(A^-)$.

Our first move is to 'improve' our formula $\varphi(\overline{x}, \bar{z}) \in \mathbb{P}_A$. As notation, for each $i < 2^n$, let $\varphi_i(\overline{x}, \bar{y}_j : j < i)$ be the restriction of φ to the smaller set of variables (we write $\bar{y}_j$ in place of the more cumbersome $\bar{y}_{s_j}$). Call an A-commitment φ *self-sufficient* if, for every $0 < i < 2^n - 1$,

$$\varphi_i(\overline{x}, \bar{y}_j : j < i) \ \vdash \ \exists \bar{y}_i \, \varphi_{i+1}(\overline{x}, \bar{y}_j : j \leq i)$$

The notion of a self-sufficient commitment is a variant on what Hrushovski and Shelah call an 'S-condition' in [HS91]. There, with Proposition 2.3(a) they prove:

Claim: For any fmac A, every $\varphi \in \mathbb{P}_A$, has a self-sufficient $\varphi^* \in \mathbb{P}_A$ extending φ. Moreover, if $f : A \to \omega_1$ witnesses that $\varphi \in \mathbb{P}_A$, then the same function f witnesses that $\varphi^* \in \mathbb{P}_A$.

Given the Claim, to verify **Splitting** we may assume that φ itself is self-sufficient. Choose an injective function $f : A \to \omega_1$ and tuples $\bar{b}_i$ from $M(f(s_i)$ for each $i < 2^n$ such that $N \models \varphi(\bar{c}_f, \bar{b}_i : i < 2^n)$, where $\bar{b}_i$ is short for $\bar{b}_{s_i}$. Getting half of the witnessing set is routine, and just amounts to adjusting the notation. Let $f_0 : A_0 \to \omega_1$ be defined as $f_0(a\hat{\ }0) = f(a)$ and $f_0(a') = f(a')$ for all $a' \in A^-$. In particular, $\bar{c}_{f_0} = \bar{c}_f$ so f_0 witnesses that $h_0(\varphi)$ is consistent. Write $\bar{c}_f$ as $c_0\hat{\ }\bar{c}^*$. The second half will require us to find an element $c' \in C \setminus \bar{c}_f$ so that $\text{tp}(c'/M(\emptyset))$ is sufficiently close to $\text{tp}(c_0/M(\emptyset))$ and then finding tuples $\langle \bar{b}'_i : 2^{n-1} \leq i < 2^n \rangle$ from the stable system. First, note that c_0 is independent from $\bar{c}^*$ over $M(\emptyset)$. Coupled with the fact that each $\bar{b}_i$ is dominated by $\{c_{f(a)} : a \in s_i\}$ over $M(\emptyset)$, there is a formula $\delta(x) \in \text{tp}(c_0/M(\emptyset))$ so that if c' is any realization of δ that is independent from $\bar{c}^*$ over $M(\emptyset)$, then

$$N \models \varphi_i(c'\bar{c}^*, \bar{b}_j : j < i) \quad \text{for all } i < 2^{n-1}$$

However, Clause (1) of our hypotheses on our stable system imply that there is some $c_\beta \in C \setminus \bar{c}_f$ that satisfies these requirements. Now, define $f_1 : A_1 \to \omega_1$ by $f_1(a\hat{\ }1) = \beta$ and $f_1(a') = f(a')$ for all $a' \in A^-$. Then, using the self-sufficiency of φ, one recursively finds tuples $\bar{b}'_j$ from $M(f_1(s_j))$ for each $2^{n-1} \leq j < 2^n$ such that

$$N \models \varphi_k(c_\beta\bar{c}^*, \langle \bar{b}_i : i < 2^{n-1} \rangle, \langle \bar{b}'_j : 2^{n-1} \leq j < k \rangle) \text{ for each } 2^{n-1} \leq k < 2^n$$

Combining these two halves yields that $f^* = f_0 \cup f_1$ witnesses that $\varphi' := h_0(\varphi) \wedge h_1(\varphi) \wedge x_{a\hat{\ }0} \neq x_{a\hat{\ }1}$ is in $\mathbb{P}_{A*a}$.

With the verification of **Splitting** in hand, Theorem 6.4.5 and hence Theorem 6.4.1 follow immediately by an application of Theorem 5.4.

Remark 6.4.6. This result does not immediately translate to the study of complete sentences of $L_{\omega_1,\omega}$. While stability notions are defined in that context ([Bal09]), the superstability hypothesis on the ambient theory here is vastly stronger than infinitary stability which concerns only the atomic models.

7 Applications II – Theories with Skolem functions

In this section we give applications of the Henkin method outlined in the previous sections to construct customized models of size continuum of theories that have Skolem functions. We first indicate how the existence of Skolem functions allows for a streamlining of our technique. Recall that if T is a complete theory that has Skolem functions, then given any model M of T, the Skolem hull of any subset $C \subseteq M$ will be an elementary substructure $N \preceq M$ in which each $b \in N$ is the interpretation of $\tau(c_1, \ldots, c_k)$ for some L-term τ and some sequence $(c_1, \ldots, c_k)$ of distinct elements of C. In particular, having such tight control obviates the need for Y-variables! More precisely, extra elements are needed to close X to a model, but the existence of Skolem functions makes their interpretations unique, and thus redundant. Within this section, we will take $Z = X = \{x_\eta : \eta \in 2^\omega\}$ as our set of variable symbols and we will construct a complete type $\Gamma(X)$ that is consistent with T. As noted above, since T admits Skolem functions, simply by taking the definable closure of any realization of Γ inside any model, $\Gamma(X)$ uniquely determines a model of T.

Thus, if T has definable Skolem functions, then the **Henkin witnesses** condition becomes vacuous. As we are only concerned with X-variables, the **Completeness** and **Splitting** are easier to verify. As usual, the **Modeling** T clause is satisfied so long as every formula describing a commitment is satisfied in a model of T. However, more care must be taken with **Omitting** Δ. In particular, our construction has to ensure that no X-instantiated L-term $t(x_{\eta_1}, \ldots, x_{\eta_n})$ (or m-tuple of terms if Δ is m-ary) realizes Δ. In practice this will be easy to ensure, so long as the 'witnessing models' each omit Δ.

It might seem that definable Skolem functions are in irreconcilable conflict with the existence of large atomic models. Indeed, if such a theory is countable, it cannot have an uncountable atomic model. Despite that, we can use the technique here to construct atomic models of size continuum by expanding the language as follows.

Definition 7.0.1. A *representation* of an $L(\Phi)$-$L_{\omega_1,\omega}$-sentence Φ is a triple $(L, T', \Delta(w))$ such that L is a countable extension of $L(\Phi)$, T is an L-theory, and $\Delta(w)$ is a 1-type such that $Mod(\Phi)$ is equal to the class of $L(\Phi)$-reducts of models of T' that omit Δ. Abusing notation somewhat, a *Skolemized representation* is a representation in which T' admits definable Skolem functions, admits elimination of quantifiers, and has a pairing function.

Applying Remark 3.3, it easy to find $(L, T', \Delta(w))$, a Skolemized representation, for an arbitrary complete $L_{\omega_1,\omega}$-sentence; choose a countable language $L' \supseteq L$, a first-order L'-theory T', and a partial type $\Delta(w)$ such that the models of Φ are precisely the L-reducts of models of T' that omit $\Delta(w)$. By expanding the language still further (but maintaining countability) we may assume $(L, T', \Delta(w))$ is a Skolemized representation. Then, if we construct a model M' of T' of size continuum that omits each of the partial types Δ_n given in the proof of Remark 3.3.2, its reduct M to L is a large atomic model of T.

7.1 Two-cardinal models

In a pair of papers, [She75b, She76], Shelah proves a celebrated two-cardinal transfer theorem. For us, it is noteworthy as this is apparently the first place where he uses the concept of asymptotic similarity. In this situation we are able to simplify by assuming Skolem functions as just discussed in the introduction.

Let T be a theory in a countable language L with a distinguished unary predicate U. A model M of T is a (κ, λ)-*model* if M has cardinality κ, but $|U(M)| = \lambda$. We are interested in constructing a $(2^{\aleph_0}, \aleph_0)$-model of T. Clearly, we will not be able to succeed for an arbitrary theory T, but we seek a sufficient condition on T for a $(2^{\aleph_0}, \aleph_0)$-model to exist.

Suppose that a countable theory T has Skolem functions. Thus, as suggested in the introduction to this section, take $X = \{x_\eta : \eta \in 2^\omega\}$ to be our distinguished set of variables, and let $\Gamma(X)$ be the **partial** type in these variables satisfying:

1. $\neg U(x_\eta)$ and $x_\eta \neq x_{\eta'}$ for distinct $\eta, \eta' \in 2^\omega$;
2. For each $k, \ell \in \omega$, for each k-ary L-term $\tau(w_1, \ldots, w_k)$ and for all pairs of ℓ-similar k-tuples $\overline{\eta} = (\eta_1, \ldots, \eta_k)$ and $\overline{\eta}' = (\eta'_1, \ldots, \eta'_k)$ we have:
$$U(\tau(x_{\eta_1}, \ldots, x_{\eta_k})) \rightarrow \left[\tau(x_{\eta_1}, \ldots, x_{\eta_k}) = \tau(x_{\eta'_1}, \ldots, x_{\eta'_k})\right]$$

The following Lemma is immediate.

Lemma 7.1.1 (Shelah,[She75b]). *Suppose that T is a countable L-theory with Skolem functions. If $T \cup \Gamma(X)$ is consistent, then T has a $(2^{\aleph_0}, \aleph_0)$-model.*

Proof. Choose a model $M \models T$ with a subset $\{c_\eta : \eta \in 2^\omega\}$ satisfying $\Gamma(X)$. It is easily checked that the Skolem hull of $\{c_\eta : \eta \in 2^\omega\}$ is a $(2^{\aleph_0}, \aleph_0)$-model of T.

But when is the type $\Gamma(X)$ consistent with T? By compactness, it suffices to show that every finite subset of $\Gamma(X)$ is consistent with T. That is, it suffices to show that every partial type $\Gamma_{\mathcal{T}}(X_F)$ is consistent with T, where $\mathcal{T}$ *is a finite set of* L*-terms* (of various arities), F is a finite subset of 2^ω, and $\Gamma_{\mathcal{T}}(X_F)$ is the finite subset of $\Gamma(X)$ that mention only terms $\tau \in \mathcal{T}$ and variables $\{x_\eta : \eta \in F\}$.

For the remainder of this discussion, fix a finite set $\mathcal{T}$ of L-terms. Note that for any finite set $F \subseteq 2^\omega$, there is a $k < \omega$ such that $\{\eta|k : \eta \in F\}$ are distinct elements of 2^k. Choose any $m \geq k$, and consider the standard fmac $2^m \subseteq 2^{<\omega}$. Let $\Gamma_{\mathcal{T}}(X_{2^m})$ be the set of X_{2^m}-instantiated formulas formed by replacing each variable symbol $x_\eta \in X_F$ by $x_{\eta|m} \in X_{(2^m)}$. As any finite tuple $\bar{c}$ from any model $M \models T$ (indeed, any L-structure) realizes $\Gamma_{\mathcal{T}}(X_F)$ if and only if it realizes $\Gamma_{\mathcal{T}}(X_{2^m})$, in order to show that $T \cup \Gamma_{\mathcal{T}}(X)$ is consistent, it suffices to prove that $T \cup \Gamma_{\mathcal{T}}(X_{2^m})$ is consistent for each of the standard fmacs 2^m.

This overview of the proof was clear to Shelah at the time he wrote [She75b], but it took him over a year to work out the combinatorics in [She76] that led to the proof of $(\aleph_\omega, \aleph_0) \to (2^{\aleph_0}, \aleph_0)$. We can now view his arguments as a slight variant on **Splitting**. Indeed, with our finite choice $\mathcal{T}$ of terms remaining fixed, choose any fmac A (and an enumeration $\langle a_i : i < n\rangle$ thereof). Suppose $M \models T$ and $\bar{c} = \langle c_a : a \in A\rangle \in M^n$ is a tuple from M realizing $\Gamma_{\mathcal{T}}(X_A)$. Choose any $a \in A$ (say $a = a_j$). We want to find an element $c^* \in M \setminus \{c_a : a \in A\}$ so that the $(n+1)$-tuple $\bar{c}\,\hat{}\,c^*$ realizes $\Gamma_{\mathcal{T}}(X_{A^{*a}})$. To obtain a sufficient condition for this, consider the equivalence relation E_n on $(M)^n$, the set of n-tuples of *distinct* elements from M given by $E_n(\bar{c}, \bar{d})$ if and only if:

> For each $\tau(\overline{w}) \in \mathcal{T}$ and corresponding subsequences $\bar{c}', \bar{d}'$ with $\lg(\overline{w}) = \lg(\bar{c}') = \lg(\bar{d}')$, **either** $M \models \neg U(\tau(\bar{c}')) \wedge \neg U(\tau(\bar{d}'))$ **or** $M \models \tau(\bar{c}') = \tau(\bar{d}')$.

It is easily verified that if M is a (κ, λ)-model, then E_n is an equivalence relation on $(M)^n$ with at most λ classes. In terms of the discussion above, given $\bar{c} \in (M)^n$, we are seeking c^* such that $E_n(\bar{c}, \bar{c}^*)$ holds, were $\bar{c}^*$ is formed by replacing c_i by c^* in $\bar{c}$.

Finally, recall that every fmac A can be constructed from $\{\langle\rangle\}$ by a sequence of $(|A|-1)$ splittings. The following Proposition is merely a restatement of Theorem 5 of [She76], noting that any equivalence relation E_n on $(M)^n$ with at most λ classes can be identified with a function $f\colon (M)^n \to \lambda$.

Proposition 7.1.2 (Shelah). *Fix any m, let $n = 2^m - 1$, and fix a sequence $\langle A_\ell : \ell \leq n\rangle$ of fmacs and a sequence $\langle a_\ell : \ell < n\rangle$ such that $A_0 = \{\langle\rangle\}$, $A_n = 2^m$, and each $A_{\ell+1} = (A_\ell)^{*a_\ell}$. If M is a (λ^{+n}, λ)-model of T, then there is a tuple $\bar{c} = \langle c_0, \cdots, c_n\rangle$ such that for every $0 < \ell \leq n$, $M \models E_\ell(\bar{c}\restriction_\ell, \bar{c}\restriction_\ell^*)$, where $\bar{c}\restriction_\ell^*$ is obtained by substituting c_n for the element of $\bar{c}_\ell$ coded by a_ℓ. In particular, for each $\ell \leq n$, $\bar{c}_\ell$ realizes $\Gamma_{\mathcal{T}}(X_{A_\ell})$.*

Given this Proposition, the following Theorem of Shelah is immediate.

Theorem 7.1.3 (Shelah, [She76]). $(\aleph_\omega, \aleph_0) \to (2^{\aleph_0}, \aleph_0)$. *Indeed, if for every n, a theory T admits a gap n-model, i.e. a $((\lambda_n)^{+n}, \lambda_n)$-model, then T admits a $(2^{\aleph_0}, \aleph_0)$-model.*

Proof. First, we may assume T has Skolem functions. Next, by Lemma 7.1.1 we need only show that $T \cup \Gamma(X)$ is consistent. Fix any finite set $\mathcal{T}$ of L-terms. By applying the Proposition for each m, we obtain the consistency of $T \cup \Gamma_{\mathcal{T}}(X_{2^m})$ for each of the standard fmacs 2^m, so we finish by compactness.

The proof of Theorem 7.1.3 is an early exemplar of the 'method of identities' which has had many applications to prove two cardinal theorems and compactness theorem in logics with generalized quantifiers. See the account in [SV06].

7.2 What is the Hanf number for an atomic model in the continuum?

Classically, a 'Hanf number' for a class of structures is the least cardinal λ such that if the class of structures has one of size λ, then it has arbitrarily large structures. For example, Morley proved that if a sentence Φ of $L_{\omega_1,\omega}$ has a model of size $\beth_{\omega_1}$, then Φ has arbitrarily large models. Here, we vary the Hanf number question by asking for the smallest cardinal λ for which the existence of a model of Φ of size λ implies the existence of a model of size continuum. Since every model of a complete $L_{\omega_1,\omega}$-sentence Φ is atomic (for a fixed expansion of the language of Φ) answering this question for a complete sentence gives the Hanf number for atomic models in the continuum.

Clearly, the value of λ can vary, depending on the size of the continuum. However, in [She99], Shelah defines (Definition 7.2.2) a cardinal $\lambda_{\omega_1}(\aleph_0)$ that is invariant under c.c.c. forcings (hence by adding enough Cohen reals, we may assume that $2^{\aleph_0} > \lambda_{\omega_1}(\aleph_0)$) and proves that if a sentence Φ of $L_{\omega_1,\omega}$ has a model of size $\lambda_{\omega_1}(\aleph_0)$, then it has a model of size $2^{\aleph_0}$.

He defines what we call (since it measures the ability to split in the sense here) a *splitting rank* for finite subsets of L-structures M in a countable language as follows:

Definition 7.2.1. For every non-empty, finite $B = \{b_0, \ldots, b_{n-1}\} \subseteq M$, we define the *splitting rank*, *sprk*(B, M), by induction on α via the following clauses:

- $\text{sprk}(B, M) \geq 0$ if $B \cap \text{acl}_M(\emptyset) = \emptyset$;
- For arbitrary α, $sprk(B, M) \geq \alpha + 1$ if and only if, for every $j < n$ and quantifier-free[15] L-formula $\varphi(w_0, \ldots, w_{n-1})$, there is $b_j^* \in (M \setminus B)$ such that
$$M \models \varphi(b_0, \ldots, b_j, \ldots, b_{n-1}) \leftrightarrow \varphi(b_0, \ldots, b_j^*, \ldots, b_{n-1})$$
and $\text{sprk}(Bb_j^*, M) \geq \alpha$; and
- For α a non-zero limit, $\text{sprk}(B, M) \geq \alpha$ if and only if $\text{sprk}(B, M) \geq \beta$ for every $\beta < \alpha$.

Then define $\text{sprk}(M) = \sup\{\text{sprk}(B, M) + 1 : B \text{ a finite subset of } M\}$ if the supremum exists, or $\text{sprk}(M) = \infty$ otherwise.

As extreme examples, suppose B is a finite subset of M satisfying $B \cap \text{acl}(\emptyset) = \emptyset$, but some $b \in B$ is algebraic over $\bar{b} = B \setminus \{b\}$. Then, if the formula $\varphi(u, \bar{b})$ witnesses the algebraicity, i.e., $M \models \varphi(b, \bar{b}) \wedge \exists^{=k} u \varphi(u, \bar{b})$, then as successive splittings of this B would require more and more distinct witnesses, we conclude that $\text{sprk}(B, M) < k$. On the other extreme, an easy induction on α shows that $\text{sprk}(B, M) \geq \alpha$ for any finite subset B of any asymptotically similar subset $\{a_\eta \colon \eta \in 2^\omega\} \subseteq M$ and any ordinal α. Thus, $\text{sprk}(M) = \infty$ whenever M contains an asymptotically similar subset.

[15] The restriction to quantifier-free formulas is inessential in our applications here, but is stated in this manner to match the usage in [She99].

Definition 7.2.2. $\lambda_{\omega_1}(\aleph_0)$ *is the least cardinal* λ *such that any structure* M *of size* λ *for any countable language necessarily has* $\operatorname{sprk}(M) \geq \omega_1$*.*

In [She99], Shelah proves that $\aleph_{\omega_1} \leq \lambda_{\omega_1}(\aleph_0) \leq \beth_{\omega_1}$ and that this cardinal is preserved under c.c.c. forcings. As the continuum can be made arbitrarily large by adding enough Cohen reals (which is a c.c.c. forcing) it is consistent that $2^{\aleph_0} > \lambda_{\omega_1}(\aleph_0)$. Despite considerable work on the problem, the question

'Does ZFC prove that $\lambda_{\omega_1}(\aleph_0) = \aleph_{\omega_1}$?'

remains open. He also gives examples of sentences Φ_α of $L_{\omega_1,\omega}$ for each $\alpha < \omega_1$ such that each Φ_α has a model M with $\operatorname{sprk}(M) = \alpha$ and no models of larger splitting rank; thus, in general, $\lambda_{\omega_1}(\aleph_0) \geq \aleph_{\omega_1}$. The main theorem of [She99] is a pleasant application of the methods developed in the previous sections:

Theorem 7.2.3 (Shelah,[She99])**.** *has a model* M *of size at least* $\lambda_{\omega_1}(\aleph_0)$*, then* Φ *has a Borel model of size continuum that contains an asymptotically similar subset* $\{c_\eta : \eta \in 2^\omega\}$*.*

Proof. Let $(L, T', \Delta(w))$ be a Skolemized representation of Φ. As T' has Skolem functions, take $Z = X = \{x_\eta : \eta \in 2^\omega\}$. We will construct a complete type $\Gamma(Z)$ that is consistent with T' and such that, if N is any model of T' and $\{c_\eta : \eta \in 2^\omega\}$ realizes $\Gamma(Z)$ in N, then the Skolem hull of $\{c_\eta : \eta \in 2^\omega\}$ will omit $\Delta(w)$.

To accomplish this, for each fmac A of $2^{<\omega}$, let $\mathbb{P}_A$ denote all instantiated formulas $\varphi(\overline{x}) = \varphi(x_a : a \in A)$ that satisfy:

> For every $\alpha < \omega_1$ there is some $\bar{b}_\alpha$ from M' that realizes φ and such that $\operatorname{sprk}(M', \bar{b}_\alpha) \geq \alpha$.

Take $\mathbb{P} = \bigcup\{\mathbb{P}_A : A \text{ an fmac}\}$ and define $\leq$ to be the usual extension relation on commitments given in Section 4.

As M' is a model of T', the structures we build will be models of T'. Also, as T' has Skolem functions, the **Henkin witnesses** conditions are trivial. More interesting verifications are:

Completeness: Fix an fmac A and an A-commitment $\varphi(x_a : a \in A) \in \mathbb{P}_A$, and choose any instantiated X_A-formula $\psi(x_a : a \in A)$. As $\varphi \in \mathbb{P}_A$, for each $\alpha < \omega_1$, choose $\bar{b}_\alpha$ from M' realizing $\varphi(\overline{x})$ with $\operatorname{sprk}(M', \bar{b}_\alpha) \geq \alpha$. There are now two cases: First, if $Y = \{\alpha < \omega_1 : M' \models \psi(\bar{b}_\alpha)\}$ is uncountable, then put $\varphi^*(\overline{x}) := \varphi \wedge \psi$. By passing to this uncountable collection, it is evident that $\varphi^* \in \mathbb{P}_A$. On the other hand, if Y is countable, then as its complement is uncountable, put $\varphi^*(\overline{x}) := \varphi \wedge \neg\psi$ and again, $\varphi^* \in \mathbb{P}_A$ and extends φ.

The verification of **Omitting** Δ is similar.

Omitting Δ: Given an fmac A and $\varphi \in \mathbb{P}_A$, choose any X_A-instantiated L-term $t(x_a : a \in A)$. As above, for each $\alpha < \omega_1$ choose a realization $\bar{b}_\alpha$ of φ in M' with $\mathrm{sprk}(M, \bar{b}_\alpha) \geq \alpha$. As M' omits $\Delta(w)$, for every α there is $\delta_\alpha(w) \in \Delta$ such that $M' \models \neg\delta_\alpha(t(\bar{b}_\alpha))$. As Δ is countable, choose a single $\delta^* \in \Delta$ such that $\{\alpha < \omega_1 : M' \models \neg\delta^*(t(\bar{b}_\alpha))\}$ is uncountable. Put $\varphi^*(\overline{x}) := \varphi \wedge \neg\delta^*(t(\overline{x}))$, which clearly extends $\varphi(\overline{x})$. By reindexing, it is evident that $\varphi^* \in \mathbb{P}_A$.

The 'shift' that occurs in the verification of **Splitting** is reminiscent of the proof of Morley's Omitting Types theorem.

Splitting: Fix any fmac A, any A-commitment $\varphi(\overline{x})$, and choose any $a \in A$. As in Definition 5.1, let $A^{*a} = A \setminus \{a\} \cup \{a\hat{\ }0, a\hat{\ }1\}$, and put $\varphi^* := \varphi(h_0(\overline{x})) \wedge \varphi(h_1(\overline{x})) \wedge x_{\delta\hat{\ }0} \neq x_{\delta\hat{\ }1}$. It suffices to show that $\varphi^* \in \mathbb{P}_{A^{*a}}$. To see this, for each $\alpha < \omega_1$, choose $\bar{b}_\alpha$ such that $M' \models \varphi(\bar{b}_\alpha)$ and $\mathrm{sprk}(M', \bar{b}_\alpha) \geq \alpha + 1$. As T' admits elimination of quantifiers, it follows from the definition of sprk that there is a 1-point extension $\bar{b}'_\alpha$ from M' extending $\bar{b}_\alpha$ that realizes φ^* with $\mathrm{sprk}(M', \bar{b}'_\alpha) \geq \alpha$. Thus, $\varphi^* \in \mathbb{P}_{A^{*a}}$.

Once all of these conditions are satisfied, it follows from Theorem 5.4 that there is a Borel model N^* of size continuum that models T' and omits $\Delta(w)$ with an asymptotically similar subset $\{c_\eta : \eta \in 2^\omega\}$. As T' has Skolem functions, the substructure $N' \preceq N^*$ generated by $\{c_\eta : \eta \in 2^\omega\}$ also models T' and omits $\Delta(w)$. Thus, as explained in the introduction to Section 7 the reduct N of N' to the original language L is a Borel model of Φ that has both size continuum and an asymptotically similar subset.

In [She99], Shelah draws an immediate Corollary from Theorem 7.2.3. Given what we have proved above, all that is required is to code the hypotheses into a suitable structure of cardinality $\lambda_{\omega_1}(\aleph_0)$.

Corollary 7.2.4 (Shelah)**.** *Let $B \subseteq 2^\omega \times 2^\omega$ be a Borel subset of the product. If B contains a $\lambda_{\omega_1}(\aleph_0)$-square (i.e., a subset $E \subseteq 2^\omega$ of size $\lambda_{\omega_1}(\aleph_0)$ such that $E \times E \subseteq B$) then there is a perfect subset E^* of the continuum with $E^* \times E^* \subseteq B$.*

Recall that classically, Morley's Omitting Types theorem states that if there is a model of power $\beth_{\omega_1}$ omitting a type, then there are Ehrenfeucht-Mostowski models that also omit the type. However, by looking more closely at the proof, the hypotheses can be weakened to: 'For every $\alpha < \omega_1$, there is a model M_α of power at least $\beth_\alpha$ that omits the type.' We note a similar analogy gives the following strengthening of Theorem 7.2.3. Specifically, to prove Theorem 7.2.5, take, for each fmac A, $\mathbb{P}_A$ to be the set of all formulas $\varphi(x_a : a \in A)$ such that for each $\alpha < \omega_1$, there is $\beta(\alpha) \geq \alpha$ and $\bar{b}_\alpha$ from $M_{\beta(\alpha)}$ realizing φ.

Theorem 7.2.5. *Suppose a sentence Φ of $L_{\omega_1,\omega}$ has a Skolemized representation (L, T', Δ). If, for every $\alpha < \omega_1$ there is a model M_α of T' that omits $\Delta(w)$ and has $\mathrm{sprk}(M_\alpha) \geq \alpha$, then there is a model N of T of size continuum that omits $\Delta(w)$ and has an asymptotically similar subset $\{c_\eta : \eta \in 2^\omega\}$.*

Theorem 7.2.5 entails the following amusing Corollary.

Corollary 7.2.6. *Let Φ be any sentence of $L_{\omega_1,\omega}$ with a Skolemized representation (L, T', Δ). If there is a ZFC-proof of the existence of a model of Φ of size continuum, then there is a Borel model of Φ with an asymptotically similar subset $\{c_\eta : \eta \in 2^\omega\}$.*

Proof. It is easily seen by induction on α that for every $\alpha < \omega_1$ there is an $L_{\omega_1,\omega}$-sentence Ψ_α in the language L' such that an L'-structure $N' \models \Psi_\alpha$ if and only if $\mathrm{sprk}(N') \geq \alpha$.

To begin the proof of the Corollary, by forcing enough Cohen reals, work in a model $\mathbb{V}[G]$ of ZFC in which $2^{\aleph_0} > \lambda_{\omega_1}(\aleph_0)$. As our forcing has the c.c.c., $\omega_1^{\mathbb{V}[G]} = \omega_1$. Working in $\mathbb{V}[G]$, choose a model $N \models \Phi$ of size continuum. Let $N' \models T'$ be an expansion of N to L' that omits Δ.

As $|N'| \geq \lambda_{\omega_1}(\aleph_0)$, $N' \models \Psi_\alpha$ for every $\alpha < \omega_1$. Thus, for each α, $\Phi \wedge \Psi_\alpha$ is formally consistent. So, working in $\mathbb{V}$, for each $\alpha < \omega_1$ an application of Karp's Completeness Theorem yields a (countable) model $M'_\alpha \models \Phi \wedge \Psi_\alpha$. Collectively, expansions of the models $\{M'_\alpha : \alpha < \omega_1\}$ satisfy the hypotheses of Theorem 7.2.5, so we finish.

Remark 7.2.7. Both Theorem 7.2.3 and Corollary 7.2.6 have analogues for atomic models. Indeed, given a countable, complete theory T, let T' be a Skolemization of T and let $\{\Delta_n\}$ be the partial types given at the end of Subsection 3.2. Let Φ be the sentence of $L_{\omega_1,\omega}$ given in Remark 3.3.2 (with respect to T'). Then, the L-reduct of any model M' of Φ will be an atomic model of T; and conversely, every atomic model M of T has an expansion to a model M' of Φ. Thus, it follows from Theorem 7.2.3 that if a countable, complete, first order theory T has an atomic model of size $\lambda_{\omega_1}(\aleph_0)$, then T has a Borel atomic model of size continuum. Similarly, the analogue of Corollary 7.2.6 is that if there is a ZFC proof of the existence of an atomic model of size continuum for a countable, complete, first order T, then there is a Borel, atomic model of T of size continuum with an asymptotically similar subset.

Remark 7.2.8. A glance at the definitions shows that having definable Skolem functions is the antithesis of dcl-triviality (see Section 6.1). In fact, the lack of non-trivial algebraic formulas directly implies that every finite subset of M has unbounded splitting rank, i.e., $\mathrm{sprk}(A, M) = \infty$ for every finite subset A of M. In fact, this 'arbitrary splitting; condition characterizes trivial-dcl. In fact, we have two proofs that theories with trivial dcl have atomic models in the continuum. The first (Subsection 6.1) took place in a extension of the given vocabulary by predicates *definable in* $L_{\omega_1,\omega}$. But the result also follows from the methods of this section using the next easy Fact and the fact that uncountable splitting rank gives a model in the continuum.

Fact 7.2.9. *The following are equivalent for an L-structure M:*

1. *M has trivial* dcl*;*
2. $\mathrm{acl}(A) = A$ *for all subsets* $A \subseteq M$*;*
3. *For every finite subset* $A \subseteq M$*,* $\mathrm{sprk}(A, M) \geq 1$*;*
4. *For every finite subset* $A \subseteq M$*,* $\mathrm{sprk}(A, M) = \infty$*;*
5. $(M, =)$ *is a sufficient pregeometry.*

References

[AFP16] N. Ackerman, C. Freer, and R. Patel. Invariant measures concentrated on countable structures. *Forum of Mathematics Sigma*, 4:e17, 59 pages, 2016.

[Bal09] John T. Baldwin. *Categoricity*. Number 51 in University Lecture Notes. American Mathematical Society, Providence, USA, 2009.

[Bal17] John T. Baldwin. The explanatory power of a new proof: Henkin's completeness proof. In M. Piazza and G. Pulcini, editors, *Philosophy of Mathematics: Truth, Existence and Explanation*, Boston Studies in the History and Philosophy of Science, page 14. Springer-Verlag, 2017. to appear: on line.

[Bay09] M. Bays. *Categoricity results for exponential maps of 1-dimensional algebraic groups & Schanuel Conjectures for Powers and the CIT*. PhD thesis, Oxford, 2009. `http://people.maths.ox.ac.uk/~bays/dist/thesis/`.

[BLS16] John T. Baldwin, C. Laskowski, and S. Shelah. Constructing many atomic models in $\aleph_1$. *Journal of Symbolic Logic*, 81:1142–1162, 2016.

[BZ11] M. Bays and B.I. Zilber. Covers of multiplicative groups of an algebraically closed field of arbitrary characteristic. *Bulletin of the London Mathematical Society*, pages 689–702, 2011.

[Hen49] Leon Henkin. The completeness of the first-order functional calculus. *Journal of Symbolic Logic*, 14:159–166, 1949.

[Hjo02] Greg Hjorth. Knight's model, its automorphism group, and characterizing the uncountable cardinals. *Journal of Mathematical Logic*, pages 113–144, 2002.

[Hjo07] Greg Hjorth. A note on counterexamples to Vaught's conjecture. *Notre Dame Journal of Formal Logic*, 2007.

[HM05] J. Hafner and P. Mancosu. The varieties of mathematical explanation. In P. Mancosu, K.F. Jorgensen, and S. Pedersen, editors, *Visualization, Explanation, and Reasoning Styles in Mathematics*, pages 251–249. Springer, 2005.

[HS91] Ehud Hrushovski and Saharon Shelah. Stability and omitting types. *Israel J Math*, 74:289–321, 1991.

[HSS09] D. Hirschfeldt, R. Shore, and T. Slaman. The atomic model theorem and type omitting. *Transactions of the American Mathematical Society*, 361:5805 – 5837, 2009.

[Kei71] H.J Keisler. *Model theory for Infinitary Logic*. North-Holland, 1971.

[Kir10] Jonathan Kirby. On quasiminimal excellent classes. *Journal of Symbolic Logic*, 75:551–564, 2010.

[KKS14] Byunghan Kim, Hyeung-Joon Kim, and Lynn Scow. Tree indiscernibilities, revisited. *Archive for Math. Logic*, 53:211–232, 2-14.

[Kni77] J.F. Knight. A complete $L_{\omega_1,\omega}$-sentence characterizing $\aleph_1$. *Journal of Symbolic Logic*, 42:151–161, 1977.

[Kue78] D. W. Kueker. Uniform theorems in infinitary logic. In A. Macintyre, L. Pacholski, and J. Paris, editors, *Logic Colloquium 77*. North Holland, 1978.

[Lac72] A.H. Lachlan. A property of stable theories. *Fundamenta Mathematicae*, 77:9–20, 1972.

[LS93] Michael C. Laskowski and Saharon Shelah. On the existence of atomic models. *Journal of Symbolic Logic*, 58:1189–1194, 1993.

[MN13] A. Montalban and A. Nies. Borel structures, a brief survey. In Noam Greenberg, Joel David Hamkins, Denis Hirschfeldt, and Russell Miller, editors, *Effective Mathematics of the Uncountable*, volume 41 of *Lecture Notes in Logic*, pages 124–134. Association of Symbolic Logic/Cambridge University Press, 2013.

[RK87] M. Resnik and D. Kushner. Explanation, independence, and realism in mathematics. *British J. Philos. Sci.*, 38:141–158, 1987.

[Ruc80] R. Rucker. *White Light*. Ace, 1980.

[She75a] S. Shelah. Categoricity in $\aleph_1$ of sentences in $L_{\omega_1,\omega}(Q)$. *Israel Journal of Mathematics*, 20:127–148, 1975. Sh index 48.

[She75b] S. Shelah. A two-cardinal theorem. *Proc American Math Soc*, 48:207–213, 1975. Sh index 37.

[She76] S. Shelah. A two-cardinal theorem and a combinatorial theorem. *Proc American Math Soc*, 62:134–136, 1976. Sh index 49.

[She78] S. Shelah. *Classification Theory and the Number of Nonisomorphic Models*. North-Holland, 1978.

[She83a] S. Shelah. Classification theory for nonelementary classes. I. the number of uncountable models of $\psi \in L_{\omega_1\omega}$ part A. *Israel Journal of Mathematics*, 46:3:212–240, 1983. Sh index 87a.

[She83b] S. Shelah. Classification theory for nonelementary classes. II. the number of uncountable models of $\psi \in L_{\omega_1\omega}$ part B. *Israel Journal of Mathematics*, 46;3:241–271, 1983. Sh index 87b.

[She99] Saharon Shelah. Borel sets with large squares. *Fundamenta Mathematica*, 159:1–50, 1999. Sh index 522.

[SV06] S. Shelah and J. Väänänen. Recursive logic frames. *Math. Logic Quart.*, 52:151–164, 2006.

[Vau61] R.L. Vaught. Denumerable models of complete theories. In *Infinitistic Methods, Proceedings of the Symposium on the Foundations of Mathematics, Warsaw, 1959*, pages 303–321. Państwowe Wydawnictwo Naukowe, Warsaw, 1961.

[Zil05] B.I. Zilber. A categoricity theorem for quasiminimal excellent classes. In *Logic and its Applications*, volume 380 of *Contemporary Mathematics*, pages 297–306. American Mathematical Society, Providence, RI, 2005.