# STABLE REGULARITY LEMMAS AND THEIR MODEL-THEORETIC FOUNDATIONS

ZE FAN

ABSTRACT. This paper studies the stable regularity lemma from both combinatorial and model-theoretic perspectives. It first presents a combinatorial proof of the stable regularity lemma for finite stable graphs (following Malliaris and Shelah [13]), utilizing the notion of $\epsilon$-excellence to construct an $\epsilon$-regular partition without irregular pairs. The necessary model-theoretic background on stability theory, including Keisler measures and rank, is then developed as a foundation for a more general approach. Building on these tools, the paper provides a model-theoretic proof of a generalized stable regularity lemma (after Malliaris and Pillay [12]) that extends the combinatorial result to arbitrary Keisler measures.

## 1. Introduction

In 1978, Szemerédi proposed his celebrated regularity lemma in [17].

**Theorem 1.1.** (Szemerédi's Regularity Lemma) *For any positive number $\epsilon$, there exists some number $N = N(\epsilon)$ and $M = M(\epsilon)$ such that for any finite graph $G = (V, R)$ with $|V| \geq N$, there exists a partition $V = V_1 \cup \cdots \cup V_k$ for some $k \leq M$ satisfying the following property:*

(1) *the partition is equitable, that is, the cardinalities of any two parts $V_i$ and $V_j$ differs by at most* 1.
(2) *all but at most $\epsilon k^2$ of the pairs $(V_i, V_j)$ are $\epsilon$-regular. That is, for any $A \subseteq V_i$ and $B \subseteq V_j$ such that $|A| \geq \epsilon |V_i|$ and $|B| \geq \epsilon |V_j|$, we have*
$$\left| \frac{|R \cap (V_i \times V_j)|}{|V_i| \cdot |V_j|} - \frac{|R \cap (A \times B)|}{|A| \cdot |B|} \right| < \epsilon.$$

This theorem poses a profound influence on fields including extremal combinatorics and theoretical computer science. The reader is referred to [10] for a wonderful survey of its various applications.

It is observed that the statement of Theorem 1.1 allows at most $\epsilon k^2$ pairs of $(V_i, V_j)$ to not be $\epsilon$-regular. Alon, Duke, Leffman, Rödl, and Yuster in [2] showed that such allowance is necessary. For a certain type of bipartite graph named *half-graph*, which has vertex sets $V = \{v_i\}_{i=1}^k$, $W = \{w_i\}_{i=1}^k$ and edge set $R = \{(v_i, w_j) : i < j\}$, non-$\epsilon$-regular pairs must exist. This observation leads to the following definition.

**Definition 1.2.** *For fixed $k \in \mathbb{N}$, a graph $G$ is said to be $k$-*edge stable *if there do not exist distinct vertices $a_1, \ldots, a_k$, $b_1, \ldots, b_k$ in $G$ such that $R(a_i, b_j)$ is true if and only if $i < j$.*

Notice that the $k$-edge stability property does not exclude just a single bipartite graph in $G$. Actually, it avoids a family of graphs being induced subgraphs of $G$, as there is no regulation on the connection state within $a_i$'s and $b_i$'s.

Following this definition, a natural question to ask is whether forbidding half-graphs from appearing in $G = (V, R)$ by stability regulations can yield an improved version of the regularity lemma that eliminates all non-$\epsilon$-regular pairs. Malliaris and Shelah give an affirmative answer to this question by proposing the Stable Regularity Lemma in [13]. Their result not only strengthened the notion of $\epsilon$-regularity by pushing the edge density between $\epsilon$-regular pairs to either less than $\epsilon$ or greater than $1 - \epsilon$, but also provided a sharper bound on the size of the partition.

**Theorem 1.3.** (Stable Regularity Lemma, [13]). *For a given $\epsilon > 0$ and $k \in \mathbb{N}$, there exists a number $N = N(\epsilon, k)$ such that for any finite $k$-edge stable graph $G$ with sufficiently large vertex set, $G$ can be partitioned into $A_1, \ldots, A_l$ for some $l \leq N$, satisfying the following properties:*

1. *the partition is equitable.*
2. *all pairs of parts $(A_i, A_j)$ are $\epsilon$-regular, and their density $d(A_i, A_j)$ is either greater than $1 - \epsilon$ or less than $\epsilon$.*
3. *the upper bound $N$ of the partition number is bounded above by $N < (4/\epsilon)^{2^{k+3}-7}$.*

This result inspired further developments in the model-theoretic study of regularity. Malliaris and Pillay in [12] obtained another version of the stable regularity lemma, which extends the notion of edge density in terms of finite counting measure to arbitrary Keisler measures, but without any bounds on partition size, as well as ignoring the partition equitability.

Earlier work by Lovász and Szegedy [11] established a version of the regularity lemma for NIP graphs with bounded VC-dimension using purely combinatorial techniques. This result predates the stable regularity lemmas discussed above, and applies to a broader class of graphs, as NIP is a weaker condition than stability. Subsequent developments combining combinatorial and model-theoretic methods have led to further refinements. For instance, Tao in his seminal paper [18] also provided a regularity lemma on algebraic hypergraphs of bounded description complexity in large finite fields, which was generalized by Chernikov and Starchenko in [5] to distal hypergraphs. Observing that stability and distality represent two opposite extreme cases of NIP (hyper)graphs, the two authors also proposed a regularity lemma for stable hypergraphs in their more recent paper [4].

This paper's main focus is on the two stable regularity lemmas by Malliaris and Shelah in [13] and by Malliaris and Pillay in [12]. In section 2, we provide an almost pure combinatorial proof for the first one. In section 3, we give a brief introduction to the necessary model-theoretic tools to understand the proof of the second stable regularity lemma. In section 4, we show the proof of this lemma, and also show that it is indeed a generalization of the first one.

## 2. Stable Regularity Lemma for Finite Graphs

In this section, we review the proof of the Stable Regularity Lemma. All graphs mentioned in this section are finite. The variables $\epsilon$ and $\zeta$ represent arbitrary positive numbers less than $1/2$.

We prove this theorem by showing that a $k$-edge stable graph $G$ can always be partitioned into parts that all have a property called $\epsilon$-excellence. This construction, by the definition presented below, guarantees that between any two parts either nearly all vertices are connected or nearly all vertices are not connected.

**Definition 2.1.** *Let $G$ be a graph, and let $\epsilon > 0$ be a fixed number. An induced subgraph $A \subseteq G$ is called* $\epsilon$-good *if for any $b \in G$, either $|\{a \in A : R(a,b)\}| < \epsilon|A|$ or $|\{a \in A : \neg R(a,b)\}| < \epsilon|A|$.*

Intuitively, if $A \subseteq G$ is $\epsilon$-good, then vertices in $A$ have a "nearly uniform opinion" on whether or not to connect to a given vertex $b \in G$. The number of outliers that disobey the opinion must be less than $\epsilon|A|$. Therefore, if for any formula $\phi$ we write $\phi^0 = \neg\phi$ and $\phi^1 = \phi$, for any $b \in G$ we can define a truth value $\mathbf{t}(b, A) \in \{0, 1\}$ such that $R(a, b)^{\mathbf{t}(b,A)}$ holds for more than $(1 - \epsilon)|A|$ vertices $a \in A$.

**Definition 2.2.** *Let $G$ be a graph, and $\epsilon > 0$ be a fixed number. An induced subgraph $A \subseteq G$ is called* $\epsilon$-excellent *if for any $\epsilon$-good subgraph $B \subseteq G$, either $|\{a \in A : \mathbf{t}(a, B) = 0\}| < \epsilon|A|$ or $|\{a \in A : \mathbf{t}(a, B) = 1\}| < \epsilon|A|$.*

An $\epsilon$-excellent subgraph $A \subseteq G$ guarantees that, every $\epsilon$-good subgraph $B \subseteq G$ has a common uniform opinion on nearly all vertices in $A$. It should also be noted that $\epsilon$-excellence implies $\epsilon$-goodness. That is because under the global assumption of $0 < \epsilon < 1/2$, the singleton $\{b\}$ for any $b \in G$ is trivially $\epsilon$-good.

In general, a large graph $G$ does not have a large $\epsilon$-excellent subset. An easy counterexample is the random graph where each edge has probability 1/2. Almost surely it has no nontrivial $\epsilon$-good sets. However, if $G$ is $k$-edge stable, then the existence of a large $\epsilon$-excellent set $A \subseteq G$ is guaranteed, where the cardinality of $A$ is bounded below by an expression related to the tree bound of $k$, which will be defined below. Since the concepts of full special tree and tree bound come from a model-theoretic context, we adopt the following model-theoretic notation.

**Notation 2.3.** *We use ${}^n2$ to denote the set of all sequences with length $n$, in which all the terms are either 0 or 1. We define ${}^{<n}2$ by ${}^{<n}2 = \bigcup_{i=0}^{n-1} {}^i2$. We use $\rho, \eta$ to represent sequences. We write $\eta|i$ to be the subsequence of $\eta$ consisting of its first $i$ terms. We write $\eta^\frown\rho$ for the sequence formed by attaching $\rho$ to the end of $\eta$. We write $\rho \trianglelefteq \eta$ if $\rho = \eta|i$ for some $i \in \mathbb{N}$.*

Now, we can define the full special tree.

**Definition 2.4.** *For $n \in \mathbb{N}$, a full special tree of height $n$ in a graph $G$ is a configuration consisting of two families of vertices in $G$, the nodes $\langle b_\rho : \rho \in {}^{<n}2\rangle$ and the leaves $\langle a_\eta : \eta \in {}^n2\rangle$, such that for any $\eta \in {}^n2$ and $\rho \in {}^{<n}2$, $\rho^\frown\langle x\rangle \trianglelefteq \eta$ if and only if $R(b_\rho, a_\eta)^x$ for $x \in \{0, 1\}$.*

A lemma in model theory implies that in a given graph $G$, the edge stability index and the maximum height of full special tree are bounded above by each other. We prove this lemma by modifying a proof in [7]. In the original proof, the definition of stability is slightly different from our definition.

**Lemma 2.5.** (see Lemma 6.7.9, p. 313, [7]). *If a graph $G$ is $k$-edge stable, then $G$ does not have any full special tree of height $2^{k+2} - 2$. If $G$ contains no full special tree of height $n$, then $G$ is $2^{n+1}$-edge stable.*

*Proof.* We first prove the second assertion by constructing a contradiction. Assume $G$ is not $2^{n+1}$-edge stable. Then there exists vertices $a_1, \dots, a_{2^{n+1}}$ and $b_1, \dots, b_{2^{n+1}}$ such that $R(a_i, b_j)$ if and only if $i > j$. We can immediately construct a full special tree of height $n$ with all the $a_i$'s as leaves and all the $b_j$'s as nodes by the following relabeling process. For a sequence $\rho \in {}^l 2$, let $\bar{\rho}$ be the natural number represented by $\rho$ if we see $\rho$ as a binary number. For example, $\overline{\langle 1, 1, 0\rangle} = 2^2 + 2 = 6$. For $\eta \in {}^n 2$, let $a_\eta = a_{2\bar{\eta}+1}$. For $\rho \in {}^i 2$, $0 \le i < n$, let $b_\rho = b_x$, where

$$x = \frac{\sum_{\eta \in {}^n 2,\ \eta | i = \rho} (2\bar{\eta} + 1)}{2^{n-k}}.$$

In this case, the $a_\eta$'s and $b_\rho$'s form a full special tree of height $n$, leading to a contradiction.

For the first assertion, we aim to prove its contrapositive. That is, if $G$ has a full special tree with height $2^{k+1} + 2$, then it is not $k$-stable. To prove this, we need to set up some more definitions. Let $T$ be a full special tree with height $k+1$. For $i \in \{0, 1\}$, we use $T_i$ to denote the $k$-tree whose nodes and leaves are the nodes $b_\rho$ and leaves $a_\eta$ of $T$ such that $\rho(0) = \eta(0) = i$. We say that a function $f : {}^{<k} 2 \to {}^{<l} 2$ is a tree map if it preserves the end-extension property of sequences. If $T$ is a $l$-tree and $N$ is a set of nodes of $H$, we say $N$ contains a $k$-tree if there exists a tree map $f : {}^{<k} 2 \to {}^{<l} 2$ such that for all $\rho \in {}^{<k} 2$, $b_\rho$ is in the set $N$. This immediately implies that there is a $k$-tree $S$ whose nodes are elements of $N$, and whose leaves are the corresponding leaves of $T$. In this case we say $N$ contains a $k$-tree $S$.

We claim that the following statement is true. For $k, l \in \mathbb{N}$ and for a full special tree $T$ with height $n + k$, if the nodes of $T$ are partitioned into two sets $N$ and $P$, then either $N$ contains a $k$-tree or $P$ contains an $l$-tree.

We prove this claim by induction on $n + k$. The base case of $n = k = 0$ is immediate. Assume $n + k > 0$, and let the vertices $b_\rho$ where $\rho \in {}^{<(n+k)} 2$ be the nodes of $T$. Without loss of generality, we also assume $b_\emptyset \in N$. For $i \in \{0, 1\}$, use $Z_i$ to denote the set of nodes of $T_i$. By induction hypothesis, for $i = 0$ or 1, either $N \cap Z_i$ contains a $(k-1)$-tree or $P \cap Z_i$ contains an $l$-tree. If at least one of $P \cap Z_0$ and $P \cap Z_1$ contains an $l$-tree, then so does $P$. If not, then both $N \cap Z_0$ and $N \cap Z_1$ contains a $(k-1)$-tree. Since $b_\emptyset \in N$, we can construct a $k$-tree in $N$ by adjoining the two $(k-1)$-trees by $b_\emptyset$. The claim is thus proved.

Back to the proof of the lemma, assume that $G$ has a full special tree with height $2^{k+1} - 2$. We will prove by induction on $k - r$ that for any $1 \le r \le n$, the following property $P_r$ holds:

> there exists vertices $a'_1, \dots, a'_{k-r}$ and $b'_1, \dots, b'_{k-r}$ of $G$, and a $(2^{r+1} - 2)$-tree $T$ in $G$, such that
> (1) for all $1 \le i, j \le k - r$, $R(a'_i, b'_j)$ if and only if $i < j$,
> (2) if $b$ is a node of $T$, then there exists a $q$ with $1 \le q \le n - r$ such that $R(a'_i, b)$ if and only if $i \le q$.
> (3) if $a$ is a leaf of $T$, then the same $q$ satisfies $R(a, b'_j)$ if and only if $j > q$.

The base case $S_k$ states that there exists a $(2^{k+1} - 2)$-tree in $G$, which is exactly our assumption. In the final case $S_1$, $T$ is a tree with height 2. Hence, it has a node $b$ and a leaf $a$ such that $R(a, b)$. Put $a$ between $a'_q$ and $a'_{q+1}$, $b$ between $b'_q$ and $b'_{q+1}$, we form a half-graph-like induced subgraph in $G$ with $2k$ vertices. That means $G$ is not $k$-edge stable.

Therefore, it remains to show the inductive step: if for $r > 1$ the property $P_r$ holds, then so does $P_{r-1}$. By the statement of $P_r$, there exists a full special tree $T$ with height $2^{r+1} - 2$. For each leaf $a$ of $T$, write $T(a)$ be the set of nodes in $T$ that have an edge with $a$. We can divide the situation into two cases.

First, assume that there exists a leaf $a$ of $T$ such that $T(a)$ contains a $(2^r - 1)$-tree. Then there is a node $b \in T(b)$ and a $(2^r - 2)$-tree $T'$ in $T(a)$ such that if we relabel all the $a'_i$ and $b'_j$ by adding 1 to their index for $i, j > q$, and setting $a'_{q+1} = a$ and $b'_{q+1} = b$, then the new sequence of $a'_i$ and $b'_j$ together with $T'$ and a new $q' = q + 1$ satisfies the statement of $P_{r-1}$.

Second, assume that for all leaves $a$ of $T$, $T(a)$ contains no $(2^r - 1)$-trees. Then let $a$ be any branch of $T_0$ and $N$ be the set of all nodes of $T_0$. The assumption implies that $N \cap T(a)$ contains no $(2^r - 1)$-tree. Thus by applying the claim to $T_0$, the set $N \backslash T(a)$ contains a $(2^r - 2)$-tree $T'$. Relabel all the $a'_i$ and $b'_j$ by adding 1 to their index for $i, j > q$, and setting $a'_{q+1} = a$ and $b'_{q+1} = b_\emptyset$, the new sequence of $a'_i$ and $b'_j$ together with $T'$ and $q$ satisfies the statement of $P_{r-1}$.

In either case $P_{r-1}$ holds. The induction is thus completed, which concludes the whole proof of the lemma. □

For a $k$-edge stable graph $G$, we can now define the *tree bound* $t = t(k)$ to be the smallest natural number such that $G$ has no full special tree of height $t$. By Lemma 2.5, $t(k)$ is well-defined and has an upper bound $t(k) \leq 2^{k+2} - 2$.

We can now prove the existence of a large $\epsilon$-excellent set in $k$-edge stable graphs.

**Proposition 2.6.** *Let $G$ be a $k$-edge stable graph, and $t = t(k)$ be its tree bound. Fix some value $\epsilon < 1/2^t$. Then, for every $A \subseteq G$ with $|A| \geq 1/\epsilon^t$, there exists an $\epsilon$-excellent set $A' \subseteq A$ with $|A'| \geq \epsilon^{t-1}|A|$.*

*Proof.* Towards contradiction, assume there exists a sufficiently large $A \subseteq G$ such that for all $A' \subseteq A$ with $|A'| \geq \epsilon^{t-1}|A|$, $A'$ is not $\epsilon$-excellent. In this case, we can define families of sets $\langle A_\eta : \eta \in {}^{<(t+1)}2\rangle$ and $\langle B_\eta : \eta \in {}^{<t}2\rangle$ by the following inductive process.

(1) For the base case, let $A_\emptyset = A$. Since $A$ is not $\epsilon$-excellent, there exists some $\epsilon$-good set $B$ that witnesses the non-excellence of $A_\emptyset$. Take $B_\emptyset = B$.
(2) For $0 < m < t$ and for $\eta \in {}^{m-1}2$, define $A_{\eta^\frown\langle i\rangle} = \{a \in A_\eta : \mathbf{t}(a, B_\eta) = i\}$ for $i \in \{0, 1\}$. Notice that both $A_{\eta^\frown\langle 0\rangle}$ and $A_{\eta^\frown\langle 1\rangle}$ have cardinality at least $\epsilon|A_\eta|$, since $B_\eta$ witnesses the non-excellence of $A_\eta$. Thus, $|A_{\eta^\frown\langle i\rangle}| \geq \epsilon|A_\eta| \geq \epsilon^m|A| \geq \epsilon^{t-1}|A|$, implying that $A_{\eta^\frown\langle i\rangle}$ is not $\epsilon$-excellent as well. Therefore, we can pick $B_{\eta^\frown\langle i\rangle}$ to be an $\epsilon$-good set that witnesses the non-excellence of $A_{\eta^\frown\langle i\rangle}$.
(3) For $\eta \in {}^t2$, we can define $A_\eta$ by the same process as in part (2). We do not define $B_\eta$ in this case, since it is not guaranteed that the sets $A_\eta$ remain non-$\epsilon$-excellence.

We now show that such a construction leads to a contradiction in the tree bound, as we can build a full special tree with height $t$ from it. To form the leaves, for each $\eta \in {}^t2$, let $a_\eta$ be any element of $A_\eta$. The set $A_\eta$ is not empty, because it has cardinality of at least $\epsilon^t|A|$, which is no less than 1 by the lower bound of $|A|$. To choose the nodes, for each $m < t$, $\rho \in {}^m2$ and $\eta \in {}^t2$ with $\rho \trianglelefteq \eta$, define the set $U_\eta = \{b \in B_\rho : R(a_\eta, b)^{1-\mathbf{t}(a_\eta, B_\rho)}\}$. We know $|U_\eta| < \epsilon|B_\rho|$ because $B_\rho$ is $\epsilon$-good. Therefore, if we further define $U_\rho = \bigcup\{U_\eta : \rho \trianglelefteq \eta\}$, then we can control the cardinality of $U_\rho$ by $|U_\rho| < 2^t\epsilon|B_\rho| < |B_\rho|$, with the last inequality coming from the fact that $\epsilon < 1/2^t$. Thus, we can pick $b_\rho$ to be any element of $B_\rho \backslash U_\rho$. The vertices $a_\eta$ and $b_\rho$ together form a $t$-tree, a contradiction. □

Base on Proposition 2.6, we may try to prove the stable regularity lemma as follows. First, pick an $\epsilon$-excellent set $A_0$ in $G$. If the remainder $G\backslash A_0$ is large enough, run the proposition once again to obtain another $\epsilon$-excellent subset $A_1 \subseteq G\backslash A_0$. Repeating this process finitely many times we can get pairwise disjoint $\epsilon$-excellent sets $A_0, \dots, A_n$, and we end this process when the remaining elements are few enough to distribute them evenly into the excellent sets without causing much trouble to their excellency. However, Proposition 2.6 has little control on the size of the excellent sets it generates, while in Theorem 1.3 we require a equitable partition. This issue can be solved by a slight modification of previous results.

**Definition 2.7.** *Let $s_0, \dots, s_{t-1}$ be a sequence of natural numbers. We call it a size sequence for $\epsilon$ when $\epsilon s_l \geq s_{l+1}$ for $l = 0, 2, \dots, t-2$, $s_{t-1}$ divides all other elements of the sequence, and $s_{t-1} > t$.*

**Proposition 2.8.** *Let $G$ be a $k$-edge stable graph, and $t = t(k)$ be its tree bound. Fix some $\epsilon < 1/2^t$.Let $s_0, \dots, s_{t-1}$ be a size sequence for $\epsilon$. Then, for any induced subgraph $A \subseteq G$ with $|A| \geq \max\{s_0, 1/\epsilon^t\}$, there exists an $\epsilon$-excellent subset $A' \subseteq A$ with $|A'| = s_l$ for some $l = 0, 1, \dots, t-1$.*

*Proof.* We use a strategy similar to the proof of Proposition 2.6. The proof will be complete if we can guarantee that during the inductive process of constructing the sets $A_\eta$'s and $B_\eta$'s, for $\eta \in {}^m 2$, the set $A_\eta$ always have size $s_m$. To handle this, in the initial stage we take $A_\emptyset$ to be any subset of $A$ with cardinality $s_0$. Then, for $0 < m \leq t$ and for $\eta \in^{m-1} 2$, by inductive hypothesis we can assume that $|A_\eta| = s_{m-1}$. Thus $A_{\eta^\frown\langle i\rangle}$ will have sizes at least $\epsilon s_{m-1} \geq s_m$ for $i \in \{0, 1\}$. If the inequality is strict, then discard some arbitrary elements in $A_{\eta^\frown\langle i\rangle}$ to make sure it has size $s_m$. □

By Proposition 2.8, we can construct from a large graph $G$ a sequence of $\epsilon$-excellent subsets whose sizes are from the terms of a fixed sequence. The next step is to refine this collection of excellent sets to make sure they form an equitable partition. To achieve this, several facts from probability theory are needed. The argument here is adopted from Section 4 of the paper [1].

**Fact 2.9.** ([16]). Consider a finite set of $N$ elements, $K$ of which possess a certain property. Define a random variable $H(s, N, K)$ as the number of elements with this property among a sample of $s$ elements drawn without replacement from the set. Then for any positive number $t$,

$$\text{Prob}\left(\frac{H(s, N, K)}{s} \geq \frac{K}{N} + t\right) \leq e^{-2t^2 s}.$$

**Proposition 2.10.** *Consider a finite set $S$ with cardinality $N$. Let $M_0, \dots, M_p$ be subsets of $S$, where $p = CN^l$ for some fixed constants $C$ and $l$. Let $r$ be some divisor of $N$. Then, for any positive integer $t$ satisfying $r \log r + \log C < 2t^2 N - rl \log N$, there exists a uniform partition (that is, a partition in which all parts have equal size) of $S$ into $r$ parts, such that* *for each piece $X$ we have*

$$\frac{|M_i \cap X|}{|X|} \leq \frac{|M_i|}{N} + t,$$

*for all $0 \leq i \leq p$.*

*Proof.* Fact 2.9 implies that

$$\text{Prob}\left(\bigvee_{i\leq p}\left(\frac{H(N/r, N, M_i)}{N/r} \geq \frac{|M_i|}{N} + t\right)\right) \leq CN^l e^{-\frac{2t^2 N}{r}}.$$

If $P = \{P_0, P_1, \ldots, P_{r-1}\}$ is a random uniform partition of $S$, then for any $0 \leq j \leq r-1$ and $0 \leq i \leq p$, the probability that $P_j$ contains at least $h$ elements in $M_i$ is $\text{Prob}(H(N/r, N, M_i) \geq h)$. Thus we have

$$\text{Prob}\left(\bigvee_{j\leq r-1}\bigvee_{i\leq p}\left(\frac{|P_j \cap M_i|}{|P_j|} \geq \frac{|M_i|}{N} + t\right)\right) \leq rCN^l e^{-\frac{2t^2 N}{r}}.$$

But if $r\log r + \log C < 2t^2 N - rl\log N$ holds, one will further have

$$\text{Prob}\left(\bigvee_{j\leq r-1}\bigvee_{i\leq p}\left(\frac{|P_j \cap M_i|}{|P_j|} \geq \frac{|M_i|}{N} + t\right)\right) < 1.$$

That means there must be some uniform partition $P$ satisfying the requirements of the proposition. □

We also need the following Proposition to proceed.

**Proposition 2.11.** *Let $k$ be a positive number, and $G$ be a $k$-edge stable graph. Then, for any subset $A \subseteq G$, $|\{\{a \in A : R(a,b)\} : b \in G\}| \leq \sum_{i=0}^{k-1}\binom{|A|}{i} \leq O(|A|^k)$.*

This proposition immediately follows from the well-know Sauer-Shelah Lemma in model theory, if we observe that the $k$-edge stability implies the graph has VC dimension at most $k-1$.

**Theorem 2.12.** (Sauer-Shelah, see Theorem II.4.10, p. 72, [15]). *Let $(V, \mathcal{F})$ be a set system having VC dimension $k$. Then, $\pi_{\mathcal{F}}(n) \leq \sum_{i=0}^{k}\binom{n}{i}$ for all positive integers $n$.*

*Proof.* We prove the theorem by induction on $n+k$. The base cases of $n=0$ and $k=0$ are trivial. For the inductive step, assume $n>0$ and $k>0$. Let $S \subseteq V$ be a arbitrary subset with cardinality $n$, and arbitrarily pick some $s \in S$. Now, for every $F \in \mathcal{F}$, define $F_S = F \cap S$. Also define $\mathcal{F}' = \{F_S \in \Pi_{\mathcal{F}}(S) : s \notin F_S, F_S \cup \{s\} \in \Pi_{\mathcal{F}}(S)\}$. It is observed that $|\Pi_{\mathcal{F}}(S)| = |\mathcal{F}'| + |\Pi_{\mathcal{F}}(S\backslash\{s\})| = |\Pi_{\mathcal{F}'}(S)| + |\Pi_{\mathcal{F}}(S\backslash\{s\})|$. Since $\mathcal{F}'$ have VC dimension at most $k-1$, by induction hypothesis we may conclude that

$$\left|\Pi_{\mathcal{F}}(S)\right| \leq \sum_{i=0}^{d-1}\binom{n}{i} + \sum_{i=0}^{d}\binom{n-1}{i} = \sum_{i=0}^{d}\binom{n}{i}.$$

The proof is thus completed. □

Now, we are ready to state the proposition concerning the construction of an equitable partition of $\epsilon$-excellent sets. This proposition is crucial for the proof of the stable regularity lemma.

**Proposition 2.13.** *Let $G$ be a finite graph with $k$-edge stable property. Fix some positive number $\zeta > \epsilon > 0$. Let $A$ be an $\epsilon$-excellent subset of $G$ with cardinality $n$. If some divisor $r$ of $n$ satisfies*

$$r \log r + \log 2 < 2(\zeta - \epsilon)^2 n - rk \log n,$$

*then there exists a uniform partition of $A$ into $r$ pieces by $A_0, A_1, \ldots, A_{r-1}$, such that each of the pieces are $\zeta$-excellent.*

*Proof.* In the context of Proposition 2.10, we define sets $M_0, M_1, \ldots, M_p$ to be in the form of $\{a \in A : R^\phi(a, b)\}$, for some $\phi \in \{0, 1\}$ and $b \in G$. By Proposition 2.11, we know $p \leq 2n^k$. Therefore, in Proposition 2.10 taking $c = 2$, $t = \zeta - \epsilon$, and $l = k$, we obtain a uniform partition of $A$ into $r$ pieces by $A_0, A_1, \ldots, A_{r-1}$, such that for all $0 \leq i \leq p$ and $0 \leq j \leq r-1$,

$$\frac{|M_i \cap A_j|}{|A_j|} \leq \frac{|M_i|}{n} + (\zeta - \epsilon).$$

That shows all the $A_j$'s are $\zeta$-good. By the fact that for any vertex $b \in G$ we always have $\mathbf{t}(b, A) = \mathbf{t}(b, A_j)$, $A_j$'s are also $\zeta$-excellent. □

It is also necessary to formalize the idea that pairs of excellent sets agree uniformly on whether or not edges exist between them. This ensures that excellent sets can serve as partitions in the stable regularity lemma.

**Proposition 2.14.** *Let $G$ be a graph and $\epsilon$ be some positive number. If subsets $A, B \subseteq G$ are both $\epsilon$-excellent, then the following statements are true.*

1. *The pair $(A, B)$ is $\epsilon$-uniform. That is, there exists a truth value $\mathbf{t} = \mathbf{t}(A, B) \in \{0, 1\}$ such that at least $(1 - \epsilon)|A|$ elements $a \in A$ and at least $(1 - \epsilon)|B|$ elements $b \in B$ satisfy $R(a, b)^{\mathbf{t}}$.*
2. *For $\zeta = (2\epsilon)^{1/2}$, the pair $(A, B)$ is $\zeta$-regular with edge density $d(A, B) > 1 - \zeta$ or $d(A, B) < \zeta$.*

*Proof.* Part (1) is immediate from the definition of excellence. To prove Part (2), for $a \in A$ we define $W_a = \{b \in B : R(a, b)^{1-\mathbf{t}}\}$ and $U = \{a \in A : |W_a| \geq \epsilon |B|\}$. It is observed that $|U| < \epsilon |A|$, and for every $a \in A \backslash U$, we have $|W_a| < \epsilon |B|$. Define $Z = \{(a, b) \in A \times B : R(a, b)^{1-\mathbf{t}}\}$. Since $W \subseteq (U \times B) \cup \bigcup_{a \notin U} \{(a, b) : b \in W_a\}$, we have the estimation

$$|W| \leq |U| \cdot |B| + |A| \cdot \max_{a \notin U} |W_a| < \epsilon |A| \cdot |B| + \epsilon |A| \cdot |B|.$$

Therefore, if $\mathbf{t} = 0$, the edge density is bounded by

$$d(A, B) = \frac{|Z|}{|A| \cdot |B|} < 2\epsilon < \zeta.$$

The last inequality holds because of the global assumption of $\epsilon < 1/2$. Similarly, if $\mathbf{t} = 1$ then we have $d(A, B) > 1 - \zeta$.

To prove $\zeta$-regularity, we arbitrarily pick $A' \subseteq A$ and $B' \subseteq B$ such that $|A'| \geq \zeta|A|$ and $|B'| \geq \zeta|B|$. Define $Z' = Z \cap (A' \times B')$. Since $Z' \subseteq (U \times B') \cup \bigcup_{a \in A' \setminus U}\{(a,b) : b \in W_a\}$, we can estimate

$$|Z'| \leq |U| \cdot |B'| + |A'| \max_{a \in A' \setminus U} |W_a| < \epsilon|A| \cdot |B'| + \epsilon|A'| \cdot |B| \leq \zeta|A'| \cdot |B'|.$$

Thus, if $\mathbf{t} = 0$, we have

$$d(A', B') = \frac{|Z'|}{|A'| \cdot |B'|} < \zeta,$$

indicating that the difference of density $|d(A,B) - d(A',B')| < \zeta$. If $\mathbf{t} = 1$, the same result can be achieved by an identical approach. Hence, the pair $(A, B)$ indeed has $\zeta$-regularity. □

Now, we are prepared to prove the stable regularity lemma. We first state and prove one of its slightly stronger forms with excellence property.

**Theorem 2.15.** *For given $\epsilon > 0$ and $k \in \mathbb{N}$, there exists a number $N = N(\epsilon, k)$ such that for any finite $k$-edge stable graph $G$ with sufficiently large vertex set, $G$ can be partitioned into $A_1, \ldots, A_l$ for some $l \leq N$, satisfying the following property:*

(1) *the partition is equitable,*
(2) *for each $i \leq l$, $A_i$ is $\epsilon$-excellent,*
(3) *all pairs of parts $(A_i, A_j)$ for $i, j \leq l$ are $\epsilon$-uniform, as defined in Proposition 2.14,*
(4) *let $t = t(k) \leq 2^{k+2} - 2$ be the tree bound, then for $\epsilon < 1/2^t$ we have $N < 4(8/\epsilon)^{t-2}$.*

*Proof.* Let $G$ be a graph with $n$ vertices, and let $t = t(k)$ be the tree bound of $k$. Without loss of generality, we assume that $\epsilon < 1/2^t$, and $\epsilon$ can be represented as a fraction with numerator 1. Define $\alpha = \epsilon/4$ and $\beta = \epsilon/3$. Also pick $q = \lceil 1/\alpha \rceil \in \mathbb{N}$, which implies $1/\alpha \leq q \leq 2/\alpha$. Let $c$ be the maximal natural number such that

$$q^{t-1}c \in \left(\frac{\alpha n}{2} - q^{t-1}, \frac{\alpha n}{2}\right].$$

We assume that $n$ is sufficiently large such that it satisfies the following three lower bounds.

(1) $\alpha^t n/2^t - 1 > t$.
(2) $\alpha^t n/2^t - 1 > M(\beta, \alpha, q^{t-1}, k)/\alpha$.
(3) $\alpha^t n/2^t - 1 > 1/\beta$.

In this case, we can define a sequence by $s_i = q^{t-1-i}c$ for $i = 0, 1, \ldots, t-1$. Observe that $s_{t-1} = c > t$. That is guaranteed by the first lower bound of $n$ by

$$c \geq \frac{\frac{\alpha n}{2} - q^{t-1}}{q^{t-1}} = \frac{\alpha n}{2q^{t-1}} - 1 \geq \frac{\alpha^t n}{2^t} - 1 > t.$$

Therefore, by applying Proposition 2.8 inductively, we can obtain a family of disjoint $\alpha$-excellent sets $\{B_j : 1 \leq j \leq j_*\}$ such that each of $B_j$ have sizes $s_l$ for some $l = 0, 1, \ldots, t-1$. Also notice that the set $B$ defined by $B = G \setminus \bigcup_{j=1}^{j_*} B_j$ has size strictly less than $s_0$.

Next, by the second lower bound of $n$, we know that the cardinality of every $B_j$ is large enough for applying Part (3) of Proposition 2.13. Thus we can further divide all the $B_j$'s, to obtain a finer partition $\{B'_i : 1 \leq i \leq i_*\} \cup \{B\}$, in which all the $B'_i$'s are $\beta$-excellent and have size $c$.

We now try to distribute elements in $B$ to the sets $B_i$'s evenly, so that after the distribution every pair of the parts are $\epsilon$-uniform. To achieve this, arbitrarily partition $B$ into $\{C_i : 1 \leq i \leq i_*\}$, where

$$|C_i| \in \left\{ \left\lfloor \frac{|B|}{i_*} \right\rfloor, \left\lfloor \frac{|B|}{i_*} \right\rfloor + 1 \right\},$$

for all $1 \leq i \leq i_*$, allowing some of the $C_i$'s being empty. We claim that for all $i$, we have $|C_i| \leq 2\beta|B_i'|$. If the claim is true, then

$$\frac{\beta|B_i'| + |C_i|}{|B_i'| + |C_i|} \leq \frac{\beta|B_i'| + 2\beta|B_i'|}{|B_i'|} = 3\beta = \epsilon,$$

implying the $\epsilon$-uniformity. To prove the claim, observe that $i_*$ is bounded by

$$\frac{n}{c} \geq i_* \geq \frac{n - (s_0 - 1)}{c} > \frac{n - s_0}{c},$$

as $|B| \leq s_0 - 1$. Thus for all $1 \leq i \leq i_*$,

$$\frac{|C_i| - 1}{|B_i'|} \leq \frac{s_0 - 1}{i_* c} < \frac{s_0 - 1}{n - s_0} < \beta,$$

where the last inequality comes from the fact that $s_0 = q^{t-1}c \leq \alpha n/2$. By the third lower bound of $n$, we have $1/c < \beta$. Thus

$$\frac{|C_i|}{|B_i'|} = \frac{|C_i| - 1}{|B_i'|} + \frac{1}{c} < \beta + \frac{1}{c} < 2\beta.$$

Therefore, if we define $A_i = C_i \cup B_i'$ for all $1 \leq i \leq i_*$, then $\{A_i : 1 \leq i \leq i_*\}$ is the equitable partition required by the theorem.

Finally, it is left to formulate a bound for $i_*$. By the choice of $c$, we have

$$\frac{\alpha n}{4q^{t-1}} \leq \frac{\alpha n}{2q^{t-1}} - 1 < c.$$

Therefore we have

$$i_* \leq \frac{n}{c} < \frac{4q^{t-1}}{\alpha} \leq \frac{4\left(\frac{2}{\alpha}\right)^{t-1}}{\alpha} = 4 \cdot \left(\frac{8}{\epsilon}\right)^{t-2}.$$

The proof is this completed. □

Theorem 1.3 is an immediate consequence of this theorem.

*Proof of Theorem 1.3.* Apply Theorem 2.15 with $k$ and $\epsilon^2/2$. The regularity can be obtain by Proposition 2.14. □

## 3. Model-Theoretic Preliminaries

The second main goal of this paper is to provide a proof of a stable regularity lemma using purely model-theoretic methods. In preparation, this section presents several foundational and standard results from model theory. Due to length considerations, proofs for most theorems and propositions are omitted; however, references to the sources containing these proofs will be explicitly indicated for the reader's convenience. We assume the reader is familiar with basic concepts in model theory, including languages, theories, models, formulas, sentences, and types.

In this paper, $L$ will always be a first order language, and $T$ will be a complete $L$-theory with only infinite models. A classical fact in model theory guarantees the existence of a sufficiently large model $M$ of $T$ with certain desired properties.

**Fact 3.1.** (Theorem 8.5, [3]) Let $\kappa$ be a regular cardinal greater than the cardinality of the set of $L$-formulas. Then there exists a model $M$ of $T$ which is $\kappa$-saturated and strongly $\kappa$-homogeneous.

Recall that $\kappa$-strong homogeneity simply means that for arbitrary subsets $A, B \subseteq M$ with cardinality less than $\kappa$ (we refer those sets as "small sets" later this text), any bijective elementary map $f : A \to B$ can be extended to an automorphism of $M$. Thus, if we pick a sufficiently large regular cardinal $\kappa_0$ that is larger than the cardinality of any specific models we consider later, and let $\mathbf{C}$ be a model satisfying Fact 3.1, then any model of $T$ with cardinality at most $\kappa_0$ is isomorphic to an elementary substructure of $\mathbf{C}$.

Let $A$ be a small subset of $\mathbf{C}$, and let $L_A$ denote the language obtained by adjoining $L$ with constant symbols for all elements of $A$. Then for a $L_A$-sentence $\phi$, we write $\models \phi$ to denote $\mathbf{C} \models \phi$. Equivalently, we know $M \models \phi$ for some model $M$ containing $A$.

For small model $M$ (with cardinality less than $\mathbf{C}$), we also follow Shelah's theory of imaginaries to construct extensions $M^{\mathrm{eq}}, L^{\mathrm{eq}}, \mathbf{C}^{\mathrm{eq}}$, and $T^{\mathrm{eq}}$. For the details of construction, please refer to Chapter 1 of Pillay, [14]. For set $A \subseteq M^{\mathrm{eq}}$, we also use $\mathrm{acl}^{\mathrm{eq}}(A)$ and $\mathrm{dcl}^{\mathrm{eq}}(A)$ to denote the algebraic closure and definable closure of $A$ over $M^{\mathrm{eq}}$.

Next, we propose the definition of stability for arbitrary $L$-formulas. It can be observed that with slight modification on indexes, this definition can be seen as a generalization of the edge stability of graphs by setting $L$ to be the language consisting only of the binary edge relation.

**Definition 3.2.** *An $L$-formula $\phi(x, y)$ is stable when there exists some $n < \omega$ such that there does not exists $a_i, b_i$ for $i \leq n$ such that $\models \phi(a_i, b_j)$ if and only if $i \leq j$.*

**Remark 3.3.** *In this definition, $x$ and $y$ can be variables or tuples of variable. Also, the formula $\phi(x, y)$ may contains additional parameters.*

If a formula $\phi(x, y)$ is stable, then its complete types (maximal consistent sets $p(x)$ consisting of formulas of the form $\phi(x, a)$ and $\neg\phi(x, a)$) have good behaviors, and the set of all its complete types $S_\phi(\mathbf{C})$ carries a well-behaved structure. To describe these phenomena in detail, we first propose the concept of definability for complete types.

**Definition 3.4.** *Let $M$ be a small model in $\mathbf{C}$, and $\delta(x, y)$ an $L$-formula without any parameters. For a complete $\delta$-type $p(x)$ over $M$, we say $p$ is $\delta$-definable if there exists an $L_M$-formula $\psi(y)$ such that for all $b \in M$, $\delta(x, b) \in p(x)$ if and only if $\models \psi(b)$. We call such $\psi$ a $\delta$-definition of $p$.*

The following important proposition guarantees the existence of $\phi$-definitions given $\phi$ is stable.

**Proposition 3.5.** (Lemma 2.2, [14]) *Let $\delta(x, y)$ be a stable $L$-formula, let $M$ be a small model, and $p(x) \in S_\delta(M)$ a complete $\delta$-type. Then, $p$ has a $\delta$-definition $\psi(y)$, which is a positive Boolean combination of formulas in the form of $\delta(a, y)$, $a \in M$.*

Next, fix a variable $x$, for any finite collection of formulas $\Delta = \Delta(x)$, we can describe the structure of $S_\Delta(\mathbf{C})$ by equipping a topology on it. For every $\Delta$-formula $\phi(x)$, if we define a set $[\phi] = \{p \in S_\Delta(\mathbf{C}) : \phi(x) \in p(x)\}$ and let all such $[\phi]$ to form a basis of a topology, then they induce a compact, Hausdorff, and totally disconnected topology on $S_\Delta(\mathbf{C})$. Recall that for any compact Hausdorff space $X$, we can define the Cantor-Bendixon rank by an inductive process. First, the rank $CB(p)$ is no less than 0 for any point $p \in X$. And $CB(p) = \alpha$ if and only if $p$ is an isolated point in the subspace $\{q \in X : CB(q) \geq \alpha\}$. Furthermore, if for some compact Hausdorff subspace $Y$ of $X$, $\alpha = \sup\{CB(p) : p \in Y\}$ is finite, then the supremum is attained and the set $Y_\alpha = \{p \in Y : CB(p) = \alpha\}$ is finite. We also define the cardinality of $Y_\alpha$ to be the CB-multiplicity of $Y$, write as $CB_{mult}(Y) = |Y_\alpha|$. The following result provides yet another demonstration of the theoretical strength of stable formulas:

**Lemma 3.6.** (Lemma 3.1, [14]) *If $\Delta = \Delta(x)$ consists only of stable formulas, then $CB(p)$ is finite for any $p \in S_\Delta(\mathbf{C})$.*

By this lemma, we can conclude that for any compact Hausdorff subspace $Y$ of $S_\Delta(\mathbf{C})$, the supremum of the Cantor-Bendixon rank on $Y$ is also finite. This leads us to the following important definition, which is very useful in a general model theory context, as well as in the proof of the stable regularity lemma in the following section.

**Definition 3.7.** *Let $\Delta = \Delta(x)$ be a finite set of stable formulas, and $\Phi(x)$ be a set of formulas with the same variable $x$, over a small subset of $\mathbf{C}$. We define the $\Delta$-rank of $\Phi(x)$, denoted $R_\Delta(\Phi(x))$ to be the Cantor-Bendixon rank of the subspace $Y = \{p \in S_\Delta(\mathbf{C}) : p(x)$ is consistent with $\Phi(x)\}$. We also define the $\Delta$-multiplicity of $\Phi(x)$ by $mult_\Delta(\Phi) = CB_{mult}(Y)$.*

It can be shown that $Y$ is always compact ad Hausdorff, thus the $\Delta$-rank must always be finite. Therefore, by applying induction on the $\Delta$-rank where $\Delta = \{\delta(x, y)\}$, one can prove the next proposition.

**Proposition 3.8.** (Lemma 1.3.7, [8]) *Let $A$ be a small set in $\mathbf{C}$, $\delta(x, y)$ be a stable formula, and $p(x) \in S_\delta(A)$. Then, there exists some global complete $\delta$-type $q(x) \in S_\delta(\mathbf{C})$ such that $p(x) \cup q(x)$ is consistent, and $q$ is $\delta$-definable over $acl^{eq}(A)$.*

By Proposition 1.3.11 in [8], Proposition 3.8 is equivalent to saying that there exists a non-forking global extension $q$ for $p$. Forking is another important theme in general model theory, which has multiple equivalent definitions. We shall propose one as follows. Recall that for some small set $A$ in $\mathbf{C}$, a sequence $(a_i : i < \alpha) \subseteq \mathbf{C}$ is *$A$-indiscernible* if for any positive integer $n$, and formula $\phi(x_1, \ldots, x_n)$ with parameters in $A$, and any increasing sequences $i_1 < \cdots < i_n$, $j_1 < \cdots < j_n$, we have $\models \phi(a_{i_1}, \ldots, a_{i_n})$ if and only if $\models \phi(a_{j_1}, \ldots, a_{j_n})$.

**Definition 3.9.** *A formula $\phi(x, a)$ divides over a small set $A$ when there exists an $A$-indiscernible sequence $(a_i : i < \omega)$ with $a_0 = a$ and $a_i \equiv_M a$ such that $\{\phi(x, a_i) : i < \omega\}$ is inconsistent. We say that a formula forks over $A$ if it implies a finite disjunction of formulas,* *each of which divides $A$. Furthermore, we say a type forks over $A$ if it implies a formula that forks over $A$.*

**Remark 3.10.** *By Corollary 1.3.13 of* [8], *if $\phi(x, y)$ and $\psi(x, b)$ are both stable formulas, and there exists some $a, b$ such that both $\phi(x, a)$ and $\psi(x, b)$ divide over some set $A$, then their disjunction $\phi(x, a) \vee \psi(x, b)$ still divides over $A$. Given this fact, it is observed that in a stable theory (a theory in which all formulas are stable) a formula forks over $A$ if and only if it divides over $A$.*

*Proof.* Dividing implying forking is trivial, thus it suffices to show that $\phi(x, a)$ forks over $A$ implies that it divides over $A$. Suppose the forking of $\phi$ is achieved by $\phi_i(x, a_i)$ for some $i = 1, 2, \ldots, n$. Since their disjunction $\bigvee^n \phi_i(x, a_i)$ still divides $A$ and $\phi(x, a)$ implies this disjunction, we can conclude that $\phi(x, a)$ divides over $A$. □

With reference to Proposition 3.8, when $A$ is a model of a stable theory, the conclusion can be sharpened: the non-forking extension is uniquely determined.

**Theorem 3.11.** (Lemma 1.4.7, [8]) *Let $M$ be a small model with stable theory, and let $\delta(x, y)$ be some formula over $M$. Then any complete $\delta$-type $p = p(x)$ has a unique non-forking global extension $q = q(x) \in S_\delta(\mathbf{C})$.*

In the stable regularity lemma discussed in the previous section, the notion of $\epsilon$-regularity was defined using an edge-counting argument. This combinatorial framework can be rigorously reformulated in measure-theoretic terms using a trivial counting measure. Leveraging model-theoretic techniques allows us to extend this combinatorial formulation naturally to finitely additive probability measures, specifically Keisler measures. Fundamental concepts and results concerning Keisler measures were first introduced in Keisler's famous paper [9]. In what follows, we present key definitions and essential tools from this theory that are necessary for our discussion.

We still work under a monster model $\mathbf{C}$, and let $M$ be a small model inside $\mathbf{C}$. A *fragment $F$ in $M$* is a small set of formulas over $M$, which contain all formulas from $L$, and is closed under connectives, quantifiers, and variable substitutions. We use $\sigma(F)$ to represent the $\sigma$-algebra of $M$, generated by all subsets that are definable in $F$. With a slight abuse of notation, we say a measure $\alpha$ *is over $F$* if it is defined on $\sigma(F)$.

**Definition 3.12.** *Given a small model $M$, a Keisler measure $\alpha$ on $M$ is a finitely additive probability measure over some fragment $F(\alpha)$ in $M$. We also say a Keisler measure is global when it is over $\mathbf{C}$.*

Now, let $\Delta$ be a finite set of stable formulas. Recall from Definition 3.7 that for any $\Delta$-type $\phi$, its $\Delta$-rank is defined by $R_\Delta(\phi) = \sup\{R_\Delta(p) : p \in [\phi]\}$. Notice that since $\Delta$ is stable, all the complete types have finite $\Delta$-rank, so the supremum in the definition is always finite and attained. Based on this notation, we can state the following proposition, whose proof is simple enough to be presented here.

**Proposition 3.13.** (Lemma 1.7, [9]) *Let $M$ be a small model, and $\alpha$ a Keisler measure on $M$ over the fragment $F = F(\alpha)$. Let $\Delta \subseteq F$ be a finite set of stable formulas, and let $\phi$ be a (stable) $\Delta$-type. Then, when $p$ runs through all the complete $\Delta$-types over $F$ such that $\alpha(\phi \cap p) > 0$, we have $\alpha(\bigcup(\phi \cap p)) = \alpha(\phi)$.*

*Proof.* We prove this proposition by induction on the $\Delta$-rank of $\phi$. The base case $R_\Delta(\phi) = 0$ is immediate: rank 0 means every point in $[\phi]$ is isolated. In a compact space, that is equivalent to $[\phi]$ is finite. Write $[\phi] = \{p_1, \dots, p_n\}$, we observe that $\bigcup_{i=1}^n (\phi \cap p_i)$ is a finite partition for $\phi$ (here we regard the types as the sets they define). By the finite additivity of $\alpha$, we get the desired equality.

For the inductive step, let $R_\Delta(\phi) = n \geq 1$. Use $F_{n-1}$ to represent the set of all $\Delta$-types over $F$ with rank at most $n-1$, then we observe that $\phi$ can be written as the union of $\phi \cap F_{n-1}$ and finitely many sets in the form $\phi \cap p$, where $p \in [\phi]$ with $R_\Delta(p) = n$. Applying induction hypothesis on $\phi_0 = \phi \wedge F_{n-1}$, we get

$$\alpha(\phi_0) = \alpha\Big(\bigcup_q (\phi_0 \cap q)\Big) = \alpha\Big(\bigcup_q (\phi \cap q)\Big),$$

where $q$ runs through all complete $\Delta$-types over $F$ with rank at most $n-1$. Adjoin this with finitely many $\phi \cap p$'s, the result gets proved. □

The concept of forking and non-forking can also be extended to Keisler measures. We propose the related definitions as follows.

**Definition 3.14.** *Let $F \subseteq G$ be two fragments, and $\phi$ a type over $F$. We say $\phi$ forks over $F$ if it forks over the set of parameters of $F$.*

*We also define the forking part of $G$ over $F$ to be the union of all complete types over $G$ that forks over $F$, written as fk$(G, F)$. If $\alpha$ and $\beta$ are two Keisler measures such that $F(\alpha) \subseteq F(\beta)$ and $\beta|_{F(\alpha)} = \alpha$, then we write fk$(\beta, \alpha) =$ fk$(F(\beta), F(\alpha))$.*

*Let $\alpha, \beta$ be Keisler measures on $M$ such that $\beta$ is an extension of $\alpha$. We say that $\beta$ is a non-forking extension of $\alpha$ when $\beta($fk$(\beta, \alpha)) = 0$.*

Analogously to Proposition 3.8, the existence of a non-forking extension for Keisler measures is guaranteed.

**Proposition 3.15.** (Theorem 1.18, [9]) *Let $M$ be a small model in $\mathbf{C}$, and let $\alpha$ a Keisler measure on $M$ over $F(\alpha)$. Then for any fragment $G \supseteq (\alpha)$, $\alpha$ has a non-forking extension to $G$.*

The last concept to be introduced in this section is that of a *countable base* of a given type $\phi$.

**Definition 3.16.** *Let $p$ be a complete type in some small model $M$ and over some fragment $F$. A fragment $G \subseteq F$ is called a countable base for $p$ if $G$ is countable, and $p|_G$ has the same $\Delta$-rank and $\Delta$-multiplicity as $p$ for any finite collection of stable formulas $\Delta \subseteq G$.*

*We also say $G$ is a countable base for a Keisler measure $\alpha$ on $M$ if $G$ is countable, $G \subseteq F(\alpha)$, and*

$$\alpha\Big(\big\{c \in M : G \text{ is a countable base for } tp(c/F(\alpha))\big\}\Big) = 1.$$

**Theorem 3.17.** (Proposition 1.20, [9]) *Let $M$ be an arbitrary small model. Every Keisler measure $\alpha$ on $M$ has a countable base.*

This theorem gives the following corollary, which will be useful in the next section.

**Corollary 3.18.** *Let $\Delta$ be a finite set of stable formulas, and $\alpha$ be a global Keisler measure. Then there exists some small model $M$ such that $\alpha|_\Delta$ (the restriction of $\alpha$ on $\Delta$-definable sets) does not fork over $M$. That is, if a $\Delta$-formula $\phi(x)$ has positive $\alpha$-measure, then it does not fork over $M$.*

*Proof.* Choose any small model $M$, and let $\beta = \alpha|_M$. Then by Theorem 3.17, $\beta$ admits a countable base $G \subseteq F(\beta) = F(M)$. Therefore, for any finite collection of stable formulas $\Delta \subseteq G$, we know that the set $X = \{c \in M : G \text{ is a countable base for } p_c := \mathrm{tp}(c/M)\}$ has $\beta$-measure 1. Now, let $\phi = \phi(x,a) \in \Delta$ be any formula with $\beta\big(\phi(x,a)\big) > 0$. Since $X$ has full measure, the intersection of $X$ and $\phi(x,a)$ has positive measure. That means there exists some $c \in M$ such that $c \models \phi(x,a)$ and $G$ is a countable base for the type $p_c := \mathrm{tp}(c/M)$. By the definition of countable base, we know $R_\Delta(p_c) = R_\Delta(p_c|_G)$.

We will now show that, $R_\Delta(p_c) = R_\Delta(p_c|_G)$ implies that $\phi(x,a)$ does not fork over $M$ by showing its contrapositive. Assume $\phi(x,a)$ forks over $M$, then by Remark 3.10 $\phi(x,a)$ divides over $M$. That means there exists $a_i \in M$ for $i < \omega$ such that $a_0 = a$, $a_i \equiv_M a$ and $\{\phi(x,a_i) : i < \omega\}$ is inconsistent. Now, take $Y = \{p \in S_\Delta(\mathbf{C}) : p_c|_G \subseteq p\}$, and for any $i < \omega$ define $Y_i = \{p \in Y : \phi(x,a_i) \in p\}$. Observe that each $Y_i$ is both open and closed in $Y$, and the intersection of all $Y_i$'s is empty. Therefore, there exists some $i$ such that the Cantor-Bendixon rank of $Y_i$ is strictly less than the rank of $Y$. That implies

$$R_\Delta(p_c) \leq R_\Delta(p_c|_G \cup \{\phi(x,a_i)\}) < R_\Delta(p_c|_G),$$

a contradiction. The corollary is thus proved. □

## 4. A Model-Theoretic Generalization

Now we are ready to state and discuss the model-theoretic version of the stable regularity lemma. Let $M$ be a saturated small model inside $\mathbf{C}$, and let $(V,W,R)$ be a definable bipartite graph. We call a finite Boolean combination of terms in the form $R(x,b)$ a $\Delta$-formula, and finite Boolean combination of terms in the form $R(a,y)$ a $\Delta^*$-formula, where $x$ and $y$ are variables while $a,b$ are parameters.

**Theorem 4.1.** (Theorem 1.1, [12]) *Let $(V,W,R)$ be a bipartite definable graph, where the edge relation is stable. Let $\mu$ and $\nu$ be Keisler measures on $V$ and $W$, respectively. Then, for any $\epsilon > 0$ one can partition $V$ into finitely many $\Delta$-definable subsets $V_1,\dots,V_m$, and $W$ into finitely many $\Delta^*$ definable subsets $W_1,\dots,W_n$, such that for each pair $(V_i,W_j)$, exactly one of the following statement holds.*

(a) *There exists a set $V_i'$ with $\mu$ measure less than $\epsilon\mu(V_i)$ such that for all $a \in V_i \backslash V_i'$, there exists some set $W_j^a \subseteq W_j$ with $\nu(W_j^a) < \epsilon\nu(W_j)$ such that $R(a,b)$ for all $b \in W_j \backslash W_j^a$. Dually, there exists a set $W_j'$ with $\nu$ measure less than $\epsilon\nu(W_j)$ such that for all $b \in W_j \backslash W_j'$, there exists some set $V_i^b \subseteq V_i$ with $\mu(V_i^b) < \epsilon\mu(V_i)$ such that $R(a,b)$ for all $a \in V_i \backslash V_i^b$.*

(b) *There exists a set $V_i'$ with $\mu$ measure less than $\epsilon\mu(V_i)$ such that for all $a \in V_i \backslash V_i'$, there exists some set $W_j^a \subseteq W_j$ with $\nu(W_j^a) < \epsilon\nu(W_j)$ such that $\neg R(a,b)$ for all $b \in W_j \backslash W_j^a$. Dually, there exists a set $W_j'$ with $\nu$ measure less than $\epsilon\nu(W_j)$ such* *that for all $b \in W_j \backslash W_j'$, there exists some set $V_i^b \subseteq V_i$ with $\mu(V_i^b) < \epsilon\mu(V_i)$ such that $\neg R(a,b)$ for all $a \in V_i \backslash V_i^b$.*

We first notice that by an argument identical to Proposition 2.14, Theorem 4.1 directly implies $\sqrt{2\epsilon}$-regularity between all pairs $(V_i, W_j)$.

To prove this theorem, we first state and prove the following important lemma. In the following discussion, by Corollary 3.18 we can always choose a small model $M$ such that the restriction of global Keisler measures $\mu|_\Delta$ and $\nu|_{\Delta^*}$ do not fork over $M$.

**Lemma 4.2.** *Let $\epsilon > 0$ be arbitrary. then in the definable bipartite graph $(V, W, R)$, we can partition $V$ into definable sets $V_1, \ldots, V_m$ such that for each $1 \leq i \leq m$ there exists a complete $\Delta$-type $p_i$ over $M$ with $\mu(p_i) > 0$, $V_i \in p_i$, and $\mu(V_i \backslash p_i) < \epsilon\mu(V_i)$ (Here we sometimes regard definable sets as the formulas that defines them).*

*Proof.* We prove this lemma by induction on the rank $R_\Delta(V)$. The base case of $R_\Delta(V) = 0$ is immediate as that means (similar to argument in Proposition 3.13) $V$ is consistent with only finitely many complete $\Delta$-types $p_1, \ldots p_k$, all of which are of rank 0. Hence, there are $\Delta$-formulas $\delta_1, \ldots, \delta_k$ such that $\delta_i$ isolates $p_i$. By taking $V_i = \{a \in V : M \models \delta_i(a)\}$, we observe that $\mu(V_i \backslash p_i) = 0 < \epsilon\mu(V_i)$, satisfies our requirement.

Next, assume the rank $R_\Delta(V) = n > 0$. In this case, $V$ is consistent with finitely many complete $\Delta$-types $p_1, \ldots, p_k$ such that $R_\Delta(p_i) = n$ for all $1 \leq i \leq k$. Writing $\mu(p_i) = \alpha_i$, we then divide the problem into three cases.

First, assume that all the $\alpha_i$'s are strictly positive. Since we have $\mu(p_i) = \inf\{\mu(\phi) : \phi \in p_i\}$, we can pick some $\phi_i \in p_i$ such that $\mu(\phi_i) \leq \alpha_1/(1-\epsilon)$. Setting $U_i = V \cap \phi_i$ and $V_i = U_i \backslash (\bigcup_{j=1}^{i-1} U_j)$, we construct disjoint definable sets $V_1, \ldots, V_k$ such that $\mu(V_i \backslash p_i) < \epsilon\mu(V_i)$. Now, let $U = \bigcup_{i=1}^k V_i$. We know that $V \backslash U$ has rank less than $n$. Apply the induction hypothesis on $V \backslash U$ and this case is proved.

Second, assume that some but not all of the $\alpha_i$'s are zero. Without loss of generality, we may assume that $\alpha_i > 0$ for $1 \leq i \leq l$, and $\alpha_i = 0$ for $l+1 \leq i \leq k$. Similar to the first case, for $1 \leq i \leq l$ we can find $V_i \in p_i$ such that $\mu(V_i \backslash p_i) < \epsilon\mu(V_i)$. Since all the $\alpha_j$'s for $j \geq l+1$ are equal to zero, there exists some $\Delta$-formula $\theta \in \bigcup_{j=l+1}^k p_j$ such that

$$\mu(\theta) < \frac{\epsilon\mu(V_1) - \mu(V_1 \backslash p_1)}{1 - \epsilon}.$$

Taking $V' = \theta$ (as the definable set),then we know $V'$ contains $\bigcup_{j=l+1}^k p_j$ and is disjoint from $\bigcup_{i=1}^l V_i$. Define $V_1' = V' \cup V_1$, then we have $\mu(V_1' \backslash p_1) < \epsilon\mu(V_1')$. Similar to the first case, we can now take $U = V_1' \cup V_2 \cup \cdots \cup V_l$ and apply induction hypothesis on $V \backslash U$. This case is also proved.

Third, assume that all of the $\alpha_i$'s are zero. By proposition 3.13, there must exist some complete $\Delta$-type $p$ such that $\mu(p) > 0$. So $R_\Delta(p) = m < n$. Let $X$ be a formula in $p$ of rank $m$, then we know $\mu(X) \geq \mu(p) > 0$. Thus by the induction hypothesis, we can partition $X$ into $X_1, \ldots, X_r$ together with complete $\Delta$-types $q_1, \ldots, q_r$ with positive measure and $q_i \subseteq X_i$, such that $\mu(X_i \backslash q_i) < \epsilon\mu(X_i)$ for all $1 \leq i \leq r$. Similar to the second case, we can find $V'$

disjoint from $X$, containing $\bigcup_{i=1}^{k} p_i$ and satisfying

$$\mu(V') < \frac{\epsilon\mu(W_1) - \mu(W_1 \backslash q_1)}{1-\epsilon}.$$

Defining $W_1 = W_1 \cup V'$, we observe that $\mu(W_1' \backslash q_1) < \epsilon\mu(W_1')$. Finally take $U = W_1' \cup W_2 \cup \cdots \cup W_r$, apply induction hypothesis on $V \backslash U$ and the whole proof is completed. $\square$

Similar to $V$, we can use this lemma to partition $W$ into $W_1, \ldots, W_n$ and find complete $\Delta^*$-types $q_1, \ldots, q_n$ such that $W_1 \in q_i$ and $\nu(W_i \backslash q_i) < \epsilon\nu(W_i)$. Now, we are finally ready to prove Theorem 4.1.

*Proof of Theorem 4.1.* We apply Lemma 4.2 twice to obtain partitions $V_1, \ldots, V_m, W_1, \ldots, W_n$ for $V$ and $W$ respectively with corresponding complete types $p_1, \ldots, p_m$ and $q_1, \ldots, q_n$. We now fix some $1 \leq i \leq m$ and $1 \leq j \leq n$. By Proposition 3.5, $p_i$ is definable with $R$-definition being some $\Delta^*$-formula $\psi(y)$. That means for all $b \in W(M)$, $\models \psi(b)$ if and only if $R(x, b) \in p_i(x)$. Thus if $p_i'$ is the unique non-forking global extension of $p_i$, then $\psi(y)$ is the $R(x, y)$-definition of $p_i'$. We divide the problem into two cases.

First, assume that $\psi(y) \in q_j$. By this assumption we know that for all $b \in W_j$ except a set of $\nu$-measure less than $\epsilon\nu(W_j)$ we have $\models \psi(b)$. Assume that $\models \psi(b)$, then we know $R(x, b) \in p_i'$, so $\neg R(x, b)$ divides (forks) over $M$. By Corollary 3.18, $\mu(\{a \in M : \neg R(a, b)\}) = \mu(p_i \cup \{\neg R(x, b)\}) = 0$. That proves the second clause in Part (a) of the statement of Theorem 4.1. The second clause of Part (a) can be proved dually by consider $\chi(x)$, the $\Delta^*$-definition of $q_j$.

Second, assume that $\psi(y) \notin q_j$, then $\neg\psi(y) \in q_j$. Using an identical argument as the first case, one can show that $(V_i, W_j)$ satisfies Part (b) of the statement of the theorem. The whole proof is thus completed. $\square$

Next, we will show that it is indeed true that Theorem 4.1 implies Theorem 1.3, but without the bound on the number of partitions and the equitable partition condition. This direction was briefly noted in Remark 1.3 of [12], and we now provide a more detailed version of the argument. To begin, we take a finite stable graph $(V, R)$. We first turn it into a bipartite graph $(V_R, V_L, R)$, by taking $V_R$ and $V_L$ as duplicates of $V$, and $R(a_R, b_L)$ if and only if $R(a, b)$ for $a_R \in V_R$, $b_L \in V_L$.

Now, we take a non-principal ultrafilter $\mathcal{U}$ of $\mathbb{N}$, and consider the constant sequence $(V_{R,n}, V_{L,n}, R_n)_{n\in\mathbb{N}}$, where $(V_{R,n}, V_{L,n}, R_n) = (V_R, V_L, R)$ for all $n \in \mathbb{N}$. We then take the ultraproduct $(V^*, W^*, R^*) = \prod_{n\to\mathcal{U}}(V_{R,n}, V_{L,n}, R_n)$. This is an infinite bipartite definable graph in a saturated model $M$ with a stable binary edge relation. We can also define Keisler measures $\mu, \nu$ on $V^*, W^*$ respectively by pushing forward the counting measures. Concretely, we define

$$\mu(\phi) = \lim_{n\to\mathcal{U}} \frac{|\{v \in V_{L,n} : V_{L,n} \models \phi\}|}{|V_{L,n}|},$$

and similarly for $\nu$. For a broader account of ultraproducts and ultrafilters in model theory, we refer the reader to Chapter 2 of [19]. Now, we can apply Theorem 4.1 on $(V^*, W^*, R^*)$. Notice that because the edges are symmetric, in the construction of partitions of $V^* = \bigcup_{i=1}^{m} V_i^*$ and $W^* = \bigcup_{j=1}^{n} W_j^*$ in Lemma 4.2 we can be sure that $n = m$ and $V_i^* = W_i^*$ for $1 \leq i \leq m$.

The third step is to pull the partition back. Notice that each $V_i^*$ is defined by some formula $\phi_i(x, b_i)$, where $b_i$ is some tuple in $W^*$. By Łoś's theorem (see Theorem 9.5.1, [7]), each $b_i$ is represented by a sequence $(b_{i,n})_{n\in\mathbb{N}}$ and for $\mathcal{U}$-almost all $n \in \mathbb{N}$ the formula $\phi_i(x, b_{i,n})$ yields a partition $V_{L,n} = \bigcup_{i=1}^m V_{L,n,i}$, where

$$V_{L,n,i} = \{a \in V_{L,n} : V_{L,n} \models \phi_i(a, b_{i,n})\},$$

and $V_{L,n,i}$ carries over the $\sqrt{2\epsilon}$-regularity. Take one of these $n$ and write $A_i = V_{L,n,i}$. This gives us a desired partition on the original finite graph $(V, R)$.

It should be remarked that, if we combine the combinatorial and model-theoretic approach, we can obtain a result better than both Theorem 1.3 and Theorem 4.1. Chernikov and Starchenko in [4] give the following result by basically using the combinatorial approach, and defining $\epsilon$-goodness and $\epsilon$-excellence with respect to arbitrary Keisler measures.

**Theorem 4.3.** Theorem 4.13, [4] *Let $(V, R)$ be a $d$-stable $k$-uniform hypergraph, and let $\epsilon \in (0, 1/2^d)$ be arbitrary. Then there exists an $R$-definable partition of $V^k$ with $\epsilon$-regularity of densities either greater than $1 - \epsilon$ or less than $\epsilon$. Also, the size of the partition is bounded by a polynomial of degree $d + 1$ in $1/\epsilon$.*

## References


[1] N. Ackerman, C. Freer, R. Patel, Stable regularity for relational structures, *arXiv:1712.09305*, 2018.

[2] N. Alon et al., The algorithmic aspects of the regularity lemma, J. Algorithms **16** (1994), no. 1, 80–109; MR1251840

[3] J. T. Baldwin, *Categoricity*, University Lecture Series, 50, Amer. Math. Soc., Providence, RI, 2009; MR2532039

[4] A. Chernikov and S. Starchenko, Definable regularity lemmas for NIP hypergraphs, Q. J. Math. **72** (2021), no. 4, 1401–1433; MR4350155

[5] A. Chernikov and S. Starchenko, Regularity lemma for distal structures, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 10, 2437–2466; MR3852184

[6] P. Erdős and A. Rényi, On the central limit theorem for samples from a finite population, Magyar Tud. Akad. Mat. Kutató Int. Közl. **4** (1959), 49–61; MR0107294

[7] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, 42, Cambridge Univ. Press, Cambridge, 1993; MR1221741

[8] F. Jahnke, D. Palacín, and K. Tent (eds.), *Lectures in Model Theory*, EMS Series of Lectures in Mathematics, European Mathematical Society, Zürich, 2018.

[9] H. J. Keisler, Measures and forking, Ann. Pure Appl. Logic **34** (1987), no. 2, 119–169; MR0890599

[10] J. Komlós and M. Simonovits, Szemerédi's regularity lemma and its applications in graph theory, in *Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993)*, 295–352, Bolyai Soc. Math. Stud., 2, János Bolyai Math. Soc., Budapest, ; MR1395865

[11] L. Lovász and B. Szegedy, Regularity partitions and the topology of graphons, in *An irregular mind*, 415–446, Bolyai Soc. Math. Stud., 21, János Bolyai Math. Soc., Budapest, 2010; MR2815610

[12] M. Malliaris and A. Pillay, The stable regularity lemma revisited, Proc. Amer. Math. Soc. **144** (2016), no. 4, 1761–1765; MR3451251

[13] M. Malliaris and S. Shelah, Regularity lemmas for stable graphs, Trans. Amer. Math. Soc. **366** (2014), no. 3, 1551–1585; MR3145742

[14] A. Pillay, *Geometric stability theory*, Oxford Logic Guides Oxford Science Publications, 32, Oxford Univ. Press, New York, 1996; MR1429864

[15] S. Shelah, *Classification Theory and the number of non-isomorphic models*, first ed. 1978, rev. ed. 1990, North-Holland.

[16] M. Skala, Hypergeometric tail inequalities: ending the insanity, *arXiv:1311.5939*, 2013.

[17] E. Szemerédi, Regular partitions of graphs, in *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, pp. 399–401, Colloq. Internat. CNRS, 260, CNRS, Paris, ; MR0540024

[18] T. C. Tao, Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets, Contrib. Discrete Math. **10** (2015), no. 1, 22–98; MR3386249

[19] Henry Towsner. *Randomness in the Limit*. Draft manuscript, 2022. Available at `https://www.sas.upenn.edu/~htowsner/RandomnessBook.pdf`.

Department of Mathematics and Computer Science, Wesleyan University, 265 Church Street, Middletown, CT 06459, USA

*Email address*: `zfan@wesleyan.edu`