

CLASSIFICATION OF BIPARTITE OBSERVABLES III: SIMPLICITY, MIXTURES AND QUANTUM ERROR CORRECTION

EDUARDO DUEÑEZ AND JOSÉ IOVINO

ABSTRACT. The paradigm of a simple unstable theory in the sense of Shelah is the theory of the random graph, the Fraïssé limit of the finite graphs. We show that its counterpart in quantum error correction is the Pauli group on countably many qubits, taken modulo phases and with commutation as its relation: the Fraïssé limit of the finite alternating spaces over the two-element field, whose form computes the syndromes of stabilizer codes. We show that simplicity theory reads as a theory of error correction: Clifford encoding circuits implement the homogeneity of the Pauli group, logical errors are characterized by forking, and the independence theorem of Kim and Pillay is the gluing of syndromes. Both structures define bipartite observables, and we show that for both, once preparations are mixed, the real-valued theory of the observable has the tree property of the second kind. The two paradigms reach it by different routes, dictated by their probes. We show that the probes of the Pauli observable are the Pauli channels, that the indicator of each syndrome sector is the difference of the responses of two uniform mixtures, and that actual mixtures witness the tree. We show that the probes of the Rado observable see only marginals, that conditions given by actual mixtures are compatible on disjoint blocks, and that the tree lives among idealized preparations.

CONTENTS

1. Introduction	2
2. Simple theories and the theorem of Kim and Pillay	4
3. The two paradigms	5
4. Stabilizer codes as a structure: a dictionary	7
5. Observables, mixtures and dividing	12
6. The theory of mixed preparations has the tree property of the second kind	15
7. The Pauli observable: syndrome sectors and the first check to fire	16
8. The Rado observable: marginals and idealized preparations	19
9. Affine logic and correlated preparations	22
10. Questions	22
Historical remarks	23
Declarations	23
References	23

Date: October 14, 2026.

2020 Mathematics Subject Classification. Primary 03C45, 81P73; Secondary 03C50, 03C66, 05C80, 81P45.

Key words and phrases. Simple theories, tree property of the second kind, forking, Kim–Pillay theorem, Fraïssé limits, Rado graph, Pauli group, stabilizer codes, quantum error correction, Pauli channels, bipartite observables, continuous logic.

1. INTRODUCTION

The paradigm of a simple unstable theory is the theory T_{ind} of the random graph, the motivating example of simple theories [She80, GIL02]. The random graph is the Rado graph, the Fraïssé limit of the finite graphs and the almost sure isomorphism type of a random network of couplings on countably many sites [Rad64, ER63]. In quantum error correction, errors and checks are Pauli strings, tensor products of Pauli matrices on the individual qubits. A Pauli string on n qubits is determined, up to a phase, by its label, the vector $(x, z) \in \mathbb{F}_2^{2n}$ that records which qubits carry a factor X and which carry a factor Z . Two Pauli strings either commute or anticommute: they anticommute exactly when $\omega(v, w) = 1$, where v and w are their labels and ω is the nondegenerate alternating form on $\mathbb{F}_2^{2n}$ given by $\omega((x, z), (x', z')) = x \cdot z' + z \cdot x'$. We identify each Pauli string with its label; the syndrome bit of an error e with respect to a check g is then $\omega(g, e)$.

In Section 3 we show that, with this form, the Pauli group on countably many qubits is the Fraïssé limit of the finite alternating spaces over the two-element field $\mathbb{F}_2$, and that its theory is simple and unstable as well, so that the Pauli group is the counterpart of the random graph in quantum error correction. This paper studies the two paradigms side by side. It translates model-theoretic simplicity theory as a theory of quantum error correction, and it shows that the paradigm, the Rado graph, and its quantum counterpart, the Pauli group, which look alike as first-order structures, behave differently once preparations are mixed.

The paper is the third of a series on bipartite observables. A bipartite observable compares a preparation of Alice with a probe of Bob and returns a number. In the first paper [DI26a] we showed that the observable is stable, in the sense of Shelah, exactly when its idealized responses are continuous; in the second [DI26b], that it is NIP exactly when they are of the first Baire class, so that the classification of Rosenthal compacta of Todorćević applies. The next dividing line of Shelah is simplicity. Stability and NIP are properties of the topological closure of a family of responses. Let us notice that simplicity is a property of a different kind: it asks whether conditions on an unknown probe can be satisfied together, and it fails when one can build a tree of conditions whose branches are consistent while the siblings exclude one another. In error correction, the unknown probe is the error, and the conditions are statements about its syndrome.

In Section 3 we prove, with the independence relation of Kim and Pillay, that the two paradigms (the Rado graph and the Pauli group) are simple and unstable. In Section 4 we translate the Fraïssé limit, forking and the independence theorem into the language of stabilizer codes: Clifford encoding circuits implement the homogeneity of the limit, the logical errors of a code are the undetectable errors whose type over the stabilizer group and the logical operators forks over the stabilizer group, and the independence theorem is the rule by which the syndromes of independent checks are prescribed separately.

In Section 5 we pass to observables. In the Rado observable, Alice and Bob each hold a site of the network and the observable reports whether the two sites are coupled; in the Pauli observable, Alice holds a check and Bob an error, and the observable reports whether the two anticommute. Alice's mixed preparations are probability measures on her labels, and her response is affine in the mixture, so that a single mixed preparation can carry unboundedly many labels. In Lemma 5.5 and Proposition 5.4 we show that Bob's probes differ: for the Rado observable they are arbitrary probability measures on sets of neighbors, and for the Pauli observable they are the Pauli channels, whose idealized pure elements are the Pauli strings of infinite support that carry the noise of a memory on infinitely many qubits.

In Section 6 we show that, for both observables and for every diagonal observable with the independence property, once preparations are mixed, the real-valued theory of the observable has the tree property of the second kind: a uniform mixture over a finite set of labels, compared with a uniform mixture over half of it, pins a probe down to that half, and two different halves exclude each other. In Sections 7 and 8 we determine, for each paradigm, whether actual preparations already witness the tree property of the second kind.

Theorem 1.1 (Main results).

- (A) *The Rado graph and the countable symplectic space over $\mathbb{F}_2$, which is the Pauli group on countably many qubits modulo phases, are simple and unstable (Theorems 3.1 and 3.3).*
- (B) *Let a stabilizer code be given, with stabilizer labels S and logical operators L . An undetectable error is a logical error exactly when its type over $S \cup L$ forks over S , and all logical errors have the same type over S , which does not fork over the empty set. Clifford encoding circuits implement automorphisms of the Fraïssé limit, and the independence theorem is the gluing of syndromes (Theorems 4.3 and 4.5, and Propositions 4.7 and 4.8). The probes of the Pauli observable are the Pauli channels, and the noise of a memory on countably many qubits is carried by idealized errors (Proposition 5.4).*
- (C) *For every diagonal observable whose pure responses have the independence property, and in particular for the Rado observable and the Pauli observable, the real-valued theory of its mixed preparations, probes and response has the tree property of the second kind, witnessed by the difference of two responses. Hence a simple diagonal observable is NIP (Theorem 6.1).*
- (D) *For the Pauli observable, the indicator of every syndrome sector is the difference of the responses of two uniform mixtures, and the tree property of the second kind is witnessed by actual preparations: the position of the first check to fire in a block, for instance the first domain wall of a bit flip in a repetition code, is a dividing condition (Proposition 7.1, Theorem 7.2 and Example 7.4).*
- (E) *For the Rado observable, and more generally for every observable whose pure responses are freely independent, such as the random bipartite graph, the conditions on a probe given by instances of a quantifier-free formula with actual mixtures as parameters are compatible along every sequence: no such sequence is uniformly k -inconsistent, for any k . Hence the tree property of the second kind in (C) is witnessed only by idealized preparations (Theorems 8.1 and 8.3, and Corollary 8.2).*

In Sections 7 and 8 we show that the mechanism behind (D) and (E) lies in the probes. In Proposition 5.4 we show that the pure probes of the Pauli observable are the Pauli strings of arbitrary support, which form a compact group, and that the pure responses of Alice are its characters, up to an affine change. Characters multiply, so averages of responses over cosets recover the indicators of syndrome sectors, as we show in Proposition 7.1; this is the familiar formula for the projector onto a syndrome sector as an average over the stabilizer group [Got97], and with it a single mixture can say “the checks before the j -th are silent and the j -th fires”. In Lemma 5.5 we show that the pure probes of the Rado observable are all sets of neighbors, which form the whole Cantor cube, and that the pure responses are its coordinates; averages of coordinates see only the marginals of a probe, and in Theorem 8.1 a gliding hump argument shows that conditions on marginals given by actual mixtures are

compatible on disjoint blocks. In the paradigm, the Rado graph, the tree of mixtures needs idealization; in its quantum counterpart, the Pauli group, the linear relations among the labels, which are the algebra of syndromes, put the tree within reach of actual preparations.

2. SIMPLE THEORIES AND THE THEOREM OF KIM AND PILLAY

The goal of this section is to fix the vocabulary. We work in a monster model $\mathbb{M}$ of a complete first-order theory T ; sets of parameters are small.

Definition 2.1. Let $\varphi(x, y)$ be a formula and let an integer $k \geq 2$ be given.

- (1) A sequence $(b_j)_{j \in \mathbb{N}}$ is *k-inconsistent* for φ if for all $j_1 < \dots < j_k$ there is no x with $\varphi(x, b_{j_1}) \wedge \dots \wedge \varphi(x, b_{j_k})$.
- (2) The formula φ has the *tree property* if, for some $k \geq 2$, there is a tree $(b_s)_{s \in \mathbb{N}^{<\mathbb{N}}}$ such that for every $\eta \in \mathbb{N}^{\mathbb{N}}$ the set $\{\varphi(x, b_{\eta \upharpoonright n}) : n \in \mathbb{N}\}$ is consistent and for every s the sequence $(b_{s \restriction j})_{j \in \mathbb{N}}$ is *k-inconsistent* for φ .
- (3) The formula φ has the *tree property of the first kind* if there is a tree $(b_s)_{s \in \mathbb{N}^{<\mathbb{N}}}$ such that for every $\eta \in \mathbb{N}^{\mathbb{N}}$ the set $\{\varphi(x, b_{\eta \upharpoonright n}) : n \in \mathbb{N}\}$ is consistent and for all incomparable $s, t \in \mathbb{N}^{<\mathbb{N}}$ the set $\{\varphi(x, b_s), \varphi(x, b_t)\}$ is inconsistent.
- (4) The formula φ has the *tree property of the second kind* (TP_2) if, for some $k \geq 2$, there is an array $(b_{ij})_{i, j \in \mathbb{N}}$ such that every row $(b_{ij})_{j \in \mathbb{N}}$ is *k-inconsistent* for φ and for every $f: \mathbb{N} \rightarrow \mathbb{N}$ the set $\{\varphi(x, b_{if(i)}) : i \in \mathbb{N}\}$ is consistent.
- (5) The theory T is *simple* if no formula has the tree property.

Every stable theory is simple, and the tree property of the second kind implies the tree property; by a theorem of Shelah, a theory is simple exactly when it has neither TP_2 nor the tree property of the first kind (see [She90, Theorem III.7.11]). A simple theory is stable exactly when it is NIP (see [She80, GIL02]), so a simple unstable theory has the independence property. To prove simplicity we shall not build trees; instead, we shall exhibit an independence relation directly.

Fact 2.2 (Kim and Pillay). Let $\perp$ be a ternary relation on small subsets of $\mathbb{M}$, written $A \perp_C B$, and assume that it satisfies the following:

- (i) invariance under automorphisms of $\mathbb{M}$;
- (ii) symmetry: the relation $A \perp_C B$ holds if and only if $B \perp_C A$ holds;
- (iii) monotonicity and transitivity: the relation $A \perp_C BD$ holds if and only if both $A \perp_C B$ and $A \perp_{CB} D$ hold;
- (iv) finite character: the relation $A \perp_C B$ holds if and only if $a \perp_C b$ holds for all finite tuples a from A and b from B ;
- (v) local character: for every finite tuple a and every set B there is $B_0 \subseteq B$ with $|B_0| \leq |T|$ and $a \perp_{B_0} B$;
- (vi) extension: for every tuple a and all sets $C \subseteq B$ there is a' with $\text{tp}(a'/C) = \text{tp}(a/C)$ and $a' \perp_C B$;
- (vii) independence theorem over models: if M is a model, $B \perp_M D$, $a_1 \perp_M B$, $a_2 \perp_M D$ and $\text{tp}(a_1/M) = \text{tp}(a_2/M)$, then there is a with $\text{tp}(a/MB) = \text{tp}(a_1/MB)$, $\text{tp}(a/MD) = \text{tp}(a_2/MD)$ and $a \perp_M BD$.

Then the theory T is simple and the relation $\perp$ coincides with nonforking independence (see [KP97, Theorem 4.2]).

3. THE TWO PARADIGMS

The goal of this section is Theorem 1.1(A): the Rado graph and the Pauli group on countably many qubits, taken modulo phases and with commutation as its relation, are simple unstable structures. In the subsections below we show that both are Fraïssé limits with quantifier elimination, and that in both the relation of Kim and Pillay is read off from the closure operator: equality for the graph, linear span for the Pauli group.

3.1. The Rado graph. The class of finite graphs has the amalgamation property, since two graphs that share an induced subgraph can be glued along it without new edges. Its Fraïssé limit is the Rado graph (Γ, E) , and a countable graph is this limit if and only if it has the *extension property*: for all finite disjoint $\Gamma_1, \Gamma_0 \subseteq \Gamma$ there are infinitely many $v \in \Gamma \setminus (\Gamma_1 \cup \Gamma_0)$ adjacent to every vertex of Γ_1 and to no vertex of Γ_0 [Fra54, Rad64]. The theory of (Γ, E) is $\aleph_0$ -categorical and eliminates quantifiers, so the type of a tuple over a set D is determined by its equalities with elements of D and its edges to D and within itself. By saturation, the monster model has the extension property for small disjoint sets Γ_1, Γ_0 .

The relation $E(x, y)$ has the independence property: for distinct $a_1, \dots, a_n$ and every set of indices among $1, \dots, n$, the extension property gives a vertex adjacent to a_i exactly when i is in the set. Hence the theory of the Rado graph is unstable.

Theorem 3.1. *The theory of the Rado graph is simple, and the relation $A \perp_C B$ holds if and only if $A \cap B \subseteq C$.*

Proof. We verify the hypotheses of Fact 2.2 for the relation $A \cap B \subseteq C$. Invariance, symmetry and finite character are immediate. For monotonicity and transitivity, the conditions $A \cap B \subseteq C$ and $A \cap D \subseteq C \cup B$ together say that every element of $A \cap (B \cup D)$ lies in C . For local character, let a finite tuple a and a set B be given; the set $B_0 = B \cap a$ is finite and contains $a \cap B$.

For extension, let a and $C \subseteq B$ be given, and write $a = (c, a_0)$ with c the entries of a in C . The type of a_0 over C prescribes the equalities and edges of a_0 within itself and to C ; by the extension property of $\mathbb{M}$, applied one entry at a time with the excluded set enlarged by $B \setminus C$, there is a'_0 disjoint from B with the same type over C . Then the tuple $a' = (c, a'_0)$ satisfies $a' \cap B \subseteq C$.

For the independence theorem, let M, B, D, a_1 and a_2 be as in Fact 2.2(vii). Since $\text{tp}(a_1/M) = \text{tp}(a_2/M)$, the two tuples have the same entries in M and the same remaining equalities; write $a_\ell = (m, a'_\ell)$ with a'_ℓ disjoint from M . Since $a_1 \perp_M B$ and $a_2 \perp_M D$, the tuple a'_1 is disjoint from B and the tuple a'_2 is disjoint from D . We look for a tuple x disjoint from $M \cup B \cup D$ with the following edges: within itself and to M as a'_1 (equivalently as a'_2); to $B \setminus M$ as a'_1 ; to $D \setminus M$ as a'_2 . These prescriptions do not conflict, because $B \cap D \subseteq M$. The extension property of $\mathbb{M}$ provides such an x , and by quantifier elimination the tuple $a = (m, x)$ is as required. $\square$

3.2. Pauli strings. A qubit is the Hilbert space $\mathbb{C}^2$, and its Pauli matrices are

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad XZ = -ZX.$$

A Pauli string on n qubits is, up to a phase, a tensor product $X^{x_1} Z^{z_1} \otimes \dots \otimes X^{x_n} Z^{z_n}$ with $x_k, z_k \in \mathbb{F}_2$, and it is labeled by the vector $(x, z) \in \mathbb{F}_2^{2n}$; products of strings correspond to sums of labels. For a label $v = (x, z)$ we write P_v for the Hermitian string $i^{x \cdot z} X^{x_1} Z^{z_1} \otimes \dots \otimes X^{x_n} Z^{z_n}$, where $x \cdot z$ is the number of indices k with $x_k = z_k = 1$; it satisfies $P_v^2 = I$, and the label

$(1, 1)$ of one qubit gives $P_{(1,1)} = iXZ = Y$. Two strings either commute or anticommute, and

$$P_{(x,z)}P_{(x',z')} = (-1)^{\omega((x,z),(x',z'))}P_{(x',z')}P_{(x,z)}, \quad \omega((x,z),(x',z')) = x \cdot z' + z \cdot x'.$$

The form ω is bilinear, alternating and nondegenerate on $\mathbb{F}_2^{2n}$. The *syndrome* of an error e with respect to checks $g_1, \dots, g_m$ is the vector $(\omega(g_1, e), \dots, \omega(g_m, e))$: measuring the checks reveals which of them anticommute with the error. On countably many qubits we let $V = \bigoplus_{k \in \mathbb{N}} \mathbb{F}_2^2$, with the same form; its elements are the Pauli strings that act nontrivially on finitely many qubits.

3.3. The Pauli group as a Fraïssé limit. In the language with $+$, 0 and a binary relation for $\omega(x, y) = 1$, substructures are $\mathbb{F}_2$ -subspaces with the restricted form. Finite alternating spaces amalgamate: if $A \subseteq B$ and $A \subseteq C$, write $B = A \oplus B'$ and $C = A \oplus C'$, and declare $\omega(B', C') = 0$ on $A \oplus B' \oplus C'$. Every finite alternating space embeds in a nondegenerate one, and by the theorem of Witt [Art57] every isometry between subspaces of a finite nondegenerate alternating space extends to the whole space. A back-and-forth argument along symplectic bases therefore shows that (V, ω) is the Fraïssé limit of the finite alternating spaces over $\mathbb{F}_2$. Its theory is $\aleph_0$ -categorical and eliminates quantifiers, so the type of a tuple a over a set D is determined by the linear relations of a over $\text{span}(D)$ and the values of ω on $\text{span}(aD)$. The algebraic closure of a set is its span.

Lemma 3.2. *If $a_1, \dots, a_n \in V$ are linearly independent and $\varepsilon_1, \dots, \varepsilon_n \in \mathbb{F}_2$, then there is $b \in V$ with $\omega(a_i, b) = \varepsilon_i$ for $i = 1, \dots, n$.*

Proof. The vectors a_i lie in a finite-dimensional subspace W spanned by part of the symplectic basis, and the form is nondegenerate on W . Hence the functionals $\omega(a_i, \cdot)$ restricted to W are linearly independent, and the linear map $b \mapsto (\omega(a_i, b))_{i \leq n}$ from W to $\mathbb{F}_2^n$ is onto. $\square$

By Lemma 3.2, the formula $\omega(x, y) = 1$ has the independence property: any set of syndromes is realized by some error. Hence the theory of (V, ω) is unstable. By saturation, the monster model $\mathbb{M}$ realizes, over each small subspace S , every finite-dimensional alternating space $S \oplus \text{span}(x_1, \dots, x_n)$; in particular, for every linear functional f on S there is b with $\omega(\cdot, b) = f$ on S .

Theorem 3.3. *The theory of (V, ω) is simple, and the relation $A \perp_C B$ holds if and only if $\text{span}(AC) \cap \text{span}(BC) = \text{span}(C)$.*

Proof. We verify the hypotheses of Fact 2.2. Invariance and symmetry are immediate. For finite character, a vector of $\text{span}(AC) \cap \text{span}(BC)$ outside $\text{span}(C)$ is a combination of finitely many elements of A , B and C .

For monotonicity and transitivity, put $Q_1 = \text{span}(AC)$, $Q_2 = \text{span}(BC)$, $Q_3 = \text{span}(BCD)$ and $Q_0 = \text{span}(C)$, so that $Q_2 \subseteq Q_3$ and $\text{span}(ABC) = Q_1 + Q_2$. It remains to be shown that $Q_1 \cap Q_3 = Q_0$ if and only if $Q_1 \cap Q_2 = Q_0$ and $(Q_1 + Q_2) \cap Q_3 = Q_2$. If $Q_1 \cap Q_3 = Q_0$, then we have $Q_1 \cap Q_2 \subseteq Q_0$; and if $u_1 + u_2 \in Q_3$ with $u_1 \in Q_1$ and $u_2 \in Q_2$, then we have $u_1 \in Q_1 \cap Q_3 = Q_0 \subseteq Q_2$, so that $u_1 + u_2 \in Q_2$. Conversely, if $u \in Q_1 \cap Q_3$, then we have $u \in (Q_1 + Q_2) \cap Q_3 = Q_2$, so that $u \in Q_1 \cap Q_2 = Q_0$.

For local character, let a be a finite tuple and B a set. The space $\text{span}(a) \cap \text{span}(B)$ is finite, so it is contained in $\text{span}(B_0)$ for some finite $B_0 \subseteq B$. If $u + w \in \text{span}(B)$ with $u \in \text{span}(a)$ and $w \in \text{span}(B_0)$, then we have $u \in \text{span}(a) \cap \text{span}(B) \subseteq \text{span}(B_0)$; hence we have $\text{span}(aB_0) \cap \text{span}(B) = \text{span}(B_0)$, which is $a \perp_{B_0} B$.

For extension, let a and $C \subseteq B$ be given. Replacing a by a maximal subtuple linearly independent over $\text{span}(C)$, we may assume that a is linearly independent over $\text{span}(C)$; the remaining entries are then determined. Choose a complement B' of $\text{span}(C)$ in $\text{span}(BC)$.

The alternating space $\text{span}(BC) \oplus \text{span}(x)$, in which x has the values of ω prescribed by $\text{tp}(a/C)$ on $\text{span}(xC)$ and $\omega(x, B') = 0$, is realized over $\text{span}(BC)$ by a tuple a' , which has the type of a over C and is linearly independent over $\text{span}(BC)$.

For the independence theorem, let M, B, D, a_1 and a_2 be as in Fact 2.2(vii). As before, we may assume that a_1 is linearly independent over M , and then so is a_2 . Since $a_1 \perp_M B$, the tuple a_1 is linearly independent over $\text{span}(MB)$, and likewise a_2 over $\text{span}(MD)$. For each entry, the functional $\omega(a_{1,k}, \cdot)$ on $\text{span}(MB)$ and the functional $\omega(a_{2,k}, \cdot)$ on $\text{span}(MD)$ agree on M , because $\text{tp}(a_1/M) = \text{tp}(a_2/M)$. Since $B \perp_M D$, we have $\text{span}(MB) \cap \text{span}(MD) = M$, so the two functionals glue to a single linear functional on $\text{span}(MB) + \text{span}(MD) = \text{span}(MBD)$. Realize, over $\text{span}(MBD)$, a tuple a linearly independent over $\text{span}(MBD)$ with these functionals and with the values of ω on a given by $\text{tp}(a_1/M)$. By quantifier elimination, the types satisfy $\text{tp}(a/MB) = \text{tp}(a_1/MB)$ and $\text{tp}(a/MD) = \text{tp}(a_2/MD)$, and the relation $a \perp_M BD$ holds by construction. $\square$

The two proofs follow the same pattern: in each structure, nonforking independence is the independence of the closure operator, equality for the graph and linear span for the Pauli group, and the independence theorem is the freedom to prescribe the remaining relations, edges or values of the form, on each side separately.

Remark 3.4. The proof of Theorem 3.3 applies over any finite field. For symmetric or arbitrary bilinear forms, the identification of the limit through symplectic bases and the theorem of Witt is no longer available in characteristic two, but the generic limit exists by free amalgamation, and the verification of Fact 2.2 goes through with the values $\omega(x, x)$ prescribed by the type.

4. STABILIZER CODES AS A STRUCTURE: A DICTIONARY

The goal of this section is to show that three notions of model theory, Fraïssé limits, forking and the independence theorem, are already at work in quantum error correction, under other names. After a brief account of stabilizer codes, each subsection states a theorem, proves it, and reads it twice: once for the physicist, who knows stabilizer codes but not model theory, and once for the logician, who knows model theory but not stabilizer codes. A single example runs through the section: the five-qubit code of Laflamme, Miquel, Paz and Zurek [LMPZ96] and of Bennett, DiVincenzo, Smolin and Wootters [BDSW96], the smallest code that corrects an arbitrary error on one qubit. Table 1 summarizes the dictionary.

4.1. Stabilizer codes in brief. We recall what the logician needs (see [Got97]). A *stabilizer group* on n qubits is a group $\mathcal{G}$ of Pauli strings that commute with each other and does not contain $-I$. Its common eigenspace for the eigenvalue 1 is the *code space*; if $\mathcal{G}$ has $n - k$ independent generators, the code space has dimension 2^k and encodes k qubits. In the labels of Section 3, the group $\mathcal{G}$ corresponds, up to signs, to an isotropic subspace $S \subseteq \mathbb{F}_2^{2n}$, that is, a subspace on which ω vanishes. An error is a Pauli string e ; measuring the generators $g_1, \dots, g_{n-k}$ of $\mathcal{G}$ on the corrupted state returns the *syndrome* $(\omega(g_1, e), \dots, \omega(g_{n-k}, e))$, and the decoder guesses the error from the syndrome. An error with zero syndrome is undetectable. It is harmless if it lies in $\mathcal{G}$, since it then acts trivially on the code space, and it is a *logical error* otherwise. On labels, the errors with zero syndrome form the *centralizer* $S^\perp = \{e \in \mathbb{F}_2^{2n} : \omega(s, e) = 0 \text{ for all } s \in S\}$, which contains S and has dimension $2n - (n - k) = n + k$; a set L of *logical operators* is a set of $2k$ labels with $S^\perp = S \oplus \text{span}(L)$, and the values $\omega(\ell, e)$ for $\ell \in L$ record which logical operator an error of $S^\perp$ applies to the encoded qubits. The *distance* of the code is the least number of qubits on which a logical

Model theory	Quantum error correction	
the structure (V, ω)	Pauli strings on countably many qubits, with commutation	Section 3
the age of (V, ω)	the commutation matrices of finite sets of Pauli strings	Proposition 4.2
ultrahomogeneity, automorphisms	Clifford circuits; encoding circuits of stabilizer codes	Theorem 4.3
algebraic closure of a set	the group generated by a set of Pauli strings	Section 3
the type of an error over a set	its syndrome, and whether it lies in the group generated	Theorem 4.5
a forking extension	the error becomes a product of known operators	Theorem 4.5
nonforking over $\emptyset$ with zero syndrome	an undetectable error that is not harmless	Corollary 4.6
forking over S of the type over $S \cup L$	a logical error	Proposition 4.7
the independence theorem	syndromes of independent checks can be prescribed independently	Proposition 4.8
idealized pure probes	Pauli strings of arbitrary support	Proposition 5.4
probes	Pauli channels	Proposition 5.4
the difference of two coset responses	the projector onto a syndrome sector	Proposition 7.1

TABLE 1. The Pauli group as a structure: the dictionary of Sections 4, 5 and 7.

error acts. A *Clifford unitary* is a unitary operator U that maps Pauli strings to Pauli strings under conjugation, up to a phase; on labels, conjugation by U is a linear map of $\mathbb{F}_2^{2n}$ that preserves ω .

Example 4.1. The five-qubit code has the stabilizer generators

$$g_1 = XZZXI, \quad g_2 = IXZZX, \quad g_3 = XIXZZ, \quad g_4 = ZXIXZ,$$

where $XZZXI$ stands for $X \otimes Z \otimes Z \otimes X \otimes I$, and it encodes one qubit, with logical operators $\bar{X} = XXXXX$ and $\bar{Z} = ZZZZZ$. Each of the fifteen errors that act as X , Y or Z on a single qubit has its own nonzero syndrome, and there are exactly fifteen nonzero syndromes of four bits; for instance, the error $XIIII$ has syndrome $(0, 0, 0, 1)$ and the error $IIZII$ has syndrome $(0, 0, 1, 0)$. The code therefore corrects every error on one qubit, and it has distance 3.

4.2. Fraïssé limits and Clifford circuits. In Proposition 4.2 and Theorem 4.3 below, we show that every finite piece of commutation structure occurs among Pauli strings, and that every symmetry between finite pieces is implemented by a Clifford circuit.

Proposition 4.2. *If Θ is a symmetric $d \times d$ matrix over $\mathbb{F}_2$ with zero diagonal, then there are Pauli strings $P_1, \dots, P_d$ on at most d qubits such that P_i and P_j anticommute exactly when $\Theta_{ij} = 1$.*

Proof. The matrix Θ defines an alternating form on $\mathbb{F}_2^d$. Let W_0 be its radical and choose a complement W of W_0 ; the form is nondegenerate on W , which therefore has a symplectic basis of r pairs, where $2r = \dim W$. Send the r pairs to the labels of X and Z on the first r qubits, and a basis of W_0 to the labels of X on $\dim W_0$ further qubits. This linear map preserves the form, and the images of the standard basis vectors of $\mathbb{F}_2^d$ are the required strings, on $r + \dim W_0 \leq d$ qubits. $\square$

Theorem 4.3. *Let W and W' be subspaces of $\mathbb{F}_2^{2n}$ and let $f: W \rightarrow W'$ be a linear bijection with $\omega(f(v), f(w)) = \omega(v, w)$ for all $v, w \in W$. Then there is a Clifford unitary U on n qubits with $UP_v U^* = \pm P_{f(v)}$ for all $v \in W$. If, moreover, the subspace W is isotropic and the vectors $v_1, \dots, v_m$ form a basis of W , then the unitary U can be chosen with $UP_{v_i} U^* = P_{f(v_i)}$ for all $i \leq m$.*

Proof. By the theorem of Witt, the map f extends to a linear bijection τ of $\mathbb{F}_2^{2n}$ that preserves ω . Every such τ is induced by conjugation by a Clifford unitary [CRSS98, DDM03]; let U be one. Since $UP_v U^*$ is a Pauli string up to a phase, Hermitian and with square I , the phase is a sign. For the second statement, let $\varepsilon_i \in \{\pm 1\}$ be the sign with $UP_{v_i} U^* = \varepsilon_i P_{f(v_i)}$. The vectors $f(v_i)$ are linearly independent, so by nondegeneracy there is $b \in \mathbb{F}_2^{2n}$ with $(-1)^{\omega(f(v_i), b)} = \varepsilon_i$ for all i . Conjugation by P_b multiplies $P_{f(v_i)}$ by $(-1)^{\omega(f(v_i), b)}$, so the Clifford unitary $P_b U$ has the required property. $\square$

Corollary 4.4. *Any two stabilizer groups on n qubits with the same number of independent generators are conjugate by a Clifford unitary. In particular, every code that encodes k qubits into n has an encoding circuit made of Clifford gates, which maps the trivial code with generators $Z_1, \dots, Z_{n-k}$ onto it.*

Proof. Let $g_1, \dots, g_m$ and $g'_1, \dots, g'_m$ be independent generators, with labels v_i and v'_i . Both families span isotropic subspaces, so the linear map $v_i \mapsto v'_i$ preserves ω , and Theorem 4.3 gives a Clifford unitary with $Ug_i U^* = g'_i$ for all i . The signs of the generators are absorbed as in the proof of Theorem 4.3. $\square$

On countably many qubits, every finite-dimensional subspace of V lies in the labels of the first n qubits for some n . Applying Theorem 4.3 there, and extending by the identity on the remaining qubits, shows that every isometry between finite-dimensional subspaces of (V, ω) extends to an automorphism of (V, ω) implemented by a Clifford circuit on finitely many qubits.

For the physicist. A Fraïssé limit is the unique countable structure that contains every finite configuration of a given kind and in which every symmetry between two finite configurations extends to a symmetry of the whole structure. Proposition 4.2 says that every finite pattern of commutation relations occurs among Pauli strings, and Theorem 4.3 says that every symmetry between finite pieces extends to a global one, implemented by a Clifford circuit. The Pauli group on countably many qubits is therefore the Fraïssé limit of finite commutation patterns, and the Clifford group is its symmetry group. The existence of encoding circuits for stabilizer codes, usually proved by an explicit construction, is the shadow of this homogeneity: by Corollary 4.4, the five-qubit code is the image of the trivial code with generators Z_1, Z_2, Z_3, Z_4 under a Clifford circuit, because the two generating sets have the same commutation pattern, namely none at all.

For the logician. The structure (V, ω) of Section 3 is the Pauli group modulo phases, and the Clifford group acts on it through the symplectic group, which is the automorphism group

of each finite piece $(\mathbb{F}_2^{2n}, \omega)$. Ultrahomogeneity, which in Section 3 came from the theorem of Witt and a back-and-forth argument, is implemented by unitary operators on a Hilbert space, the gates of a quantum computer. Stabilizer codes are finite isotropic subspaces, and Corollary 4.4 is ultrahomogeneity for isotropic subspaces, with the signs of the generators as additional data handled by Lemma 3.2.

4.3. Forking and logical errors. In a simple theory, forking is the notion of an extension of a type that adds information. The next theorem computes it for the Pauli group, using the description of nonforking independence in Theorem 3.3.

Theorem 4.5. *Let e be an element of a monster model $\mathbb{M}$ of the theory of (V, ω) , and let $D \subseteq B$ be small subsets of $\mathbb{M}$. Then the type $\text{tp}(e/B)$ forks over D if and only if $e \in \text{span}(B) \setminus \text{span}(D)$.*

Proof. By Theorem 3.3, the type $\text{tp}(e/B)$ does not fork over D exactly when $\text{span}(eD) \cap \text{span}(B) = \text{span}(D)$. If $e \in \text{span}(D)$, then we have $\text{span}(eD) = \text{span}(D)$, and the equality holds. If $e \notin \text{span}(D)$, then we have $\text{span}(eD) = \text{span}(D) \oplus \{0, e\}$, and the equality fails exactly when some vector $e + d$ with $d \in \text{span}(D)$ lies in $\text{span}(B)$, that is, exactly when $e \in \text{span}(B)$. $\square$

By Theorem 4.5, commutation data never fork: the type of e over B records the values $\omega(b, e)$ for $b \in B$, and these values can be prescribed freely without forking. Forking happens only when e becomes a product of the operators in B . In a stabilizer code this distinguishes the two kinds of undetectable error. The *weight* of an error is the number of qubits on which it acts.

Corollary 4.6. *Let $\mathcal{G}$ be a stabilizer group with set of labels S , and let e be a nonzero label with zero syndrome. Then the type $\text{tp}(e/S)$ forks over $\emptyset$ if and only if e is harmless, that is, $e \in S$. Consequently, the distance of the code is the least weight of a nonzero error with zero syndrome whose type over S does not fork over $\emptyset$.*

Proof. Apply Theorem 4.5 with $D = \emptyset$ and $B = S$; the span of S is S itself, and $\text{span}(\emptyset) = \{0\}$. The second statement is the definition of distance, since a logical error is a nonzero label with zero syndrome outside S . $\square$

In Proposition 4.7 below, we show that the stabilizer group alone does not tell logical errors apart, since they all have the same type over S , and that adding the logical operators to the known operators pins each of them down and characterizes them.

Proposition 4.7. *Let $\mathcal{G}$ be a stabilizer group with set of labels S , let L be a set of logical operators, and let e be a label with zero syndrome. Then the following hold:*

- (a) *the label e is a logical error if and only if $\text{tp}(e/S \cup L)$ forks over S ;*
- (b) *all logical errors have the same type over S , and this type does not fork over $\emptyset$.*

Proof. (a) Since the syndrome of e is zero, the label e lies in $S^\perp = \text{span}(S \cup L)$. By Theorem 4.5 with $D = S$ and $B = S \cup L$, the type $\text{tp}(e/S \cup L)$ forks over S if and only if $e \in \text{span}(S \cup L) \setminus S = S^\perp \setminus S$, which is the set of logical errors.

(b) By quantifier elimination, the type over S of a label $e \notin S$ is determined by its linear relations over S , of which there are none, and by the values of ω on $\text{span}(eS)$, which are $\omega(s, e) = 0$ and $\omega(e, e) = 0$ together with the values on S . Hence all labels of $S^\perp \setminus S$ have the same type over S , and it does not fork over $\emptyset$ by Corollary 4.6. $\square$

In the five-qubit code, the errors $g_1 g_2 = XYIYX$ (up to a phase) and $\bar{X} = XXXXX$ both have zero syndrome. The first lies in the stabilizer group, so its type over S forks; it is

harmless. The second does not, so its type does not fork; it flips the encoded qubit, and it acts on five qubits. The least weight of an error of the second kind is 3, attained for instance by $\bar{X}g_1 = IYYIX$ up to a phase. With $L = \{\bar{X}, \bar{Z}\}$, the logical errors $\bar{X}$, $\bar{Z}$ and $\bar{X}g_1$ have the same type over S , and their types over $S \cup L$ are those of the elements $\bar{X}$, $\bar{Z}$ and $\bar{X}g_1$ of $\text{span}(S \cup L)$.

For the physicist. Forking is the model-theoretic notion of an observation that pins something down, as opposed to one that merely narrows it. In the Pauli group, Theorem 4.5 shows that syndromes only narrow: on infinitely many qubits, no finite set of checks determines an error, however many bits it reveals. An error is pinned down, in the sense of forking, exactly when it becomes a product of known operators. In a code, by Corollary 4.6, this is the difference between a harmless undetectable error, which is a product of stabilizers and forks, and a logical error, which does not fork. The distance of the code, its single most important parameter, is the least weight of a nonforking undetectable error. Proposition 4.7 completes the picture: the stabilizer group alone sees all logical errors alike, and a logical error is pinned down, in the sense of forking, once the logical operators are added to the known operators.

For the logician. The theory of (V, ω) has a vector space pregeometry for algebraic closure, and Theorem 4.5 is the usual statement that in such a theory forking of a single element is algebraicity. In finite dimension, the information carried by a syndrome is measured by counting: if $s_1, \dots, s_m$ are linearly independent labels in $\mathbb{F}_2^{2n}$, then for each syndrome $\sigma \in \mathbb{F}_2^m$ exactly 2^{2n-m} labels e satisfy $\omega(s_i, e) = \sigma_i$ for all i , since the map $e \mapsto (\omega(s_i, e))_{i \leq m}$ is linear and onto. This count measures codimension, which forking does not see. The logical errors of a code are the elements of the centralizer of S outside S , and by Proposition 4.7 they share one type over S . Its nonforking extensions to $S \cup L$ are realized in (V, ω) by labels outside $\text{span}(S \cup L)$ that commute with S , for instance by errors on qubits that the code does not touch; the logical errors realize its forking extensions, which become algebraic once the logical operators are added.

4.4. The independence theorem and the gluing of syndromes. The independence theorem of Kim and Pillay says that two types that do not fork can be amalgamated. In the Pauli group it becomes a statement about syndromes.

Proposition 4.8. *Let C , B and D be subsets of $\mathbb{F}_2^{2n}$ with $\text{span}(BC) \cap \text{span}(DC) = \text{span}(C)$, and let ξ_1 and ξ_2 be linear functionals on $\text{span}(BC)$ and on $\text{span}(DC)$ that agree on $\text{span}(C)$. Then there is a label $e \in \mathbb{F}_2^{2n}$ with $\omega(s, e) = \xi_1(s)$ for $s \in \text{span}(BC)$ and $\omega(s, e) = \xi_2(s)$ for $s \in \text{span}(DC)$.*

Proof. Since ξ_1 and ξ_2 agree on the intersection of their domains, there is a linear functional ξ on $\text{span}(BC) + \text{span}(DC)$ extending both. Extend ξ to $\mathbb{F}_2^{2n}$. Because ω is nondegenerate, every linear functional on $\mathbb{F}_2^{2n}$ has the form $\omega(\cdot, e)$ for a unique e . $\square$

In the five-qubit code, take $C = \emptyset$, $B = \{g_1, g_2\}$ and $D = \{g_3, g_4\}$. The spans of B and D meet only in 0, so any two bits of syndrome on g_1, g_2 and any two on g_3, g_4 are realized together by one error; by Example 4.1, even by an error on a single qubit, for instance $YIIII$ for the syndrome $(1, 0 \mid 1, 1)$. If D is replaced by $\{g_1g_2\}$, the hypothesis fails, and so does the conclusion: the syndrome of g_1g_2 is the sum of those of g_1 and g_2 .

For the physicist. The independence theorem is the rule that lets partial information be combined: two sets of checks that are independent given their common part impose no constraint on each other, so their syndromes can be prescribed separately and are realized by a single error. The point of the theorem of Kim and Pillay is that this gluing rule, together

with the other properties of Fact 2.2, characterizes the simple theories [KP97]. Section 6 shows that the rule has no counterpart once preparations are mixed: for two halves of a block that meet in a quarter of it, the conditions “the probe answers 1 on this half and 0 on the rest of the block” and “on that half” are each consistent but cannot hold together, and infinitely many such conditions produce the tree property of the second kind.

For the logician. Proposition 4.8 is the independence theorem of Theorem 3.3 in finite dimension, over an arbitrary closed set rather than over a model: the gluing of functionals that agree on $\text{span}(MB) \cap \text{span}(MD)$ is the whole argument. Its physical content is that the syndromes of a set of checks are determined by a linear functional on the span of the checks, so that the consistency of prescribed syndromes is a matter of linear algebra on that span; the surjectivity of the syndrome map of linearly independent checks is Lemma 3.2.

5. OBSERVABLES, MIXTURES AND DIVIDING

The goal of this section is to transport Definition 2.1 to the responses of an observable, in the form used in the second paper, and to introduce the observables of the two paradigms. In Proposition 5.4 and Lemma 5.5 we show that their probes differ: the probes of the Pauli observable are the Pauli channels, and its idealized pure probes are the Pauli strings of infinite support; the pure probes of the Rado observable are all sets of neighbors.

5.1. Diagonal observables. Let A and B be countable sets and let $\varphi: A \times B \rightarrow \{0, 1\}$ be a function. The *diagonal observable* of φ is the operator

$$C_\varphi = \sum_{a \in A, b \in B} \varphi(a, b) |a\rangle\langle a| \otimes |b\rangle\langle b| \quad \text{on } \ell^2(A) \otimes \ell^2(B).$$

In plain words, Alice’s register holds a label a , Bob’s register holds a label b , and the observable reads $\varphi(a, b)$. A density operator ρ of Alice, that is, a positive operator of trace one on $\ell^2(A)$ that describes her preparation, enters only through its diagonal, which is a probability measure μ on A : the preparation draws a at random with law μ . The slice of C_φ by ρ , the operator on Bob’s space obtained by averaging C_φ over Alice’s preparation, is the diagonal operator of multiplication by

$$b \mapsto \sum_{a \in A} \mu(a) \varphi(a, b).$$

The probe algebra of the second paper is the unital C^* -algebra generated by these slices; it is the commutative C^* -algebra of bounded functions on B generated by the constants and the indicators $b \mapsto \varphi(a, b)$, $a \in A$.

Lemma 5.1. *Let $K_\varphi \subseteq \{0, 1\}^A$ be the closure of the set of functions $\varphi(\cdot, b)$, $b \in B$. The probe algebra of C_φ is isomorphic to $C(K_\varphi)$, the probe space is the space $\mathcal{P}(K_\varphi)$ of Radon probability measures on K_φ , and the response of the mixture μ is the affine continuous function*

$$r_\mu(\nu) = \int_{K_\varphi} \sum_{a \in A} \mu(a) \lambda(a) d\nu(\lambda) \quad (\nu \in \mathcal{P}(K_\varphi)).$$

Proof. The unital C^* -algebra generated by commuting projections is the algebra of continuous functions on the Stone space of the Boolean algebra they generate. The Boolean algebra generated by the sets $\{b : \varphi(a, b) = 1\}$ has as Stone space the closure of the image of B under $b \mapsto \varphi(\cdot, b)$, which is K_φ . States of $C(K_\varphi)$ are Radon probability measures by the Riesz representation theorem, and the formula for r_μ is the evaluation of the state ν at the slice. $\square$

The pure probes are the point masses δ_λ , $\lambda \in K_\varphi$; a pure probe is an *idealized label* of Bob, recording which of Alice's labels it answers with 1. Actual labels $b \in B$ are dense among them, since K_φ is the closure of the set of functions $\varphi(\cdot, b)$. For a label b , we write $r_\mu(b)$ for the response at the pure probe of $\varphi(\cdot, b)$.

5.2. The structure of an observable. To speak of the theory of an observable, we regard it as a structure of the real-valued logic of the second paper. For a diagonal observable C_φ , let $\mathcal{S}_\varphi$ be the two-sorted structure whose sorts are Alice's mixtures, the probability measures on A with the metric of total variation, and Bob's probes, the space $\mathcal{P}(K_\varphi)$ with a compatible metric for its compact metrizable weak* topology, and whose only predicate is the response $R(\mu, \nu) = r_\mu(\nu)$, with values in $[0, 1]$. The predicate is 1-Lipschitz in μ and continuous on the compact sort of probes, hence uniformly continuous. We say that C_φ is *simple* if the real-valued theory of $\mathcal{S}_\varphi$ is simple in the sense of Ben Yaacov [BY03]. In an elementary extension of $\mathcal{S}_\varphi$, the new elements of the sort of preparations are idealized preparations of Alice, and the tree properties of a formula are witnessed there. When a tree property is witnessed in $\mathcal{S}_\varphi$ itself, in the sense of Definition 5.2 below, we say that actual preparations witness it.

5.3. Dividing for families of functions. The responses of the second paper form a family of real functions on the probe space, and the independence property was defined for such a family. We do the same for the tree properties. The conditions are inequalities with a gap, which is what makes them survive in elementary extensions of the structure $\mathcal{S}_\varphi$.

Definition 5.2. Let Y be a set and $\mathcal{R}$ a family of real functions on Y .

- (1) A sequence $(r_j)_{j \in \mathbb{N}}$ in $\mathcal{R}$ is *uniformly k -inconsistent* if there are reals $\alpha < \beta$ such that each r_j takes a value $\geq \beta$ somewhere on Y , and for all $j_1 < \dots < j_k$ no $y \in Y$ satisfies $r_{j_\ell}(y) > \alpha$ for $\ell = 1, \dots, k$.
- (2) The family $\mathcal{R}$ has the *tree property of the second kind* (TP₂) if there are $k \geq 2$, reals $\alpha < \beta$ and an array $(r_{ij})_{i,j \in \mathbb{N}}$ in $\mathcal{R}$ such that every row is uniformly k -inconsistent with the constants α and β , and for every $f: \mathbb{N} \rightarrow \mathbb{N}$ and every n there is $y \in Y$ with $r_{if(i)}(y) \geq \beta$ for all $i < n$.

The tree property of Definition 2.1, its variant of the first kind, and TP₂ all require infinite sequences of instances that are k -inconsistent: the children of a node, or a row. When these notions are transported to a family of functions as in Definition 5.2, those sequences become uniformly k -inconsistent. A family without uniformly k -inconsistent sequences therefore does not have the tree property of the second kind. For the response of an observable, Definition 5.2 is the tree property of the second kind of the atomic formula, in the real-valued logic of the second paper; the gap between α and β is what the definitions of dividing in real-valued logic require [BY03].

5.4. The Pauli observable. The *Pauli observable* is the diagonal observable of ω on $V \times V$: Alice and Bob each hold a Pauli label, and the observable reports whether the two Pauli strings anticommute. We read Alice's label as a check and Bob's label as an error, so that the observable reads one syndrome bit. For a mixture μ , the response at an error e is the probability that a check drawn with law μ anticommutes with e . By Lemma 3.2, the observable has the independence property already on pure preparations.

Lemma 5.3. *For the Pauli observable, the space K_ω is the group V^* of all linear functionals $V \rightarrow \mathbb{F}_2$.*

Proof. Each $\omega(\cdot, e)$ is linear, and the linear functionals form a closed subset of $\{0, 1\}^V$. Conversely, a basic neighborhood of a linear functional λ prescribes the values of λ on finitely

many vectors, which may be taken linearly independent, and Lemma 3.2 gives e with $\omega(\cdot, e)$ in the neighborhood. $\square$

The functionals of V^* have a direct physical meaning. Let $\hat{V} = \prod_{k \in \mathbb{N}} \mathbb{F}_2^2$ be the group of labels of Pauli strings of arbitrary support, with the product topology; it contains V as the strings of finite support. For $a \in V$ and $u \in \hat{V}$, the value $\omega(a, u)$ is defined by the formula of Section 3, which is a finite sum because a has finite support. For p with $0 < p \leq 1$, let ν_p be the *depolarizing noise* of rate p : the product measure on $\hat{V}$ under which the labels on distinct qubits are independent, each equal to $(0, 0)$ with probability $1 - p$ and to each of $(1, 0)$, $(0, 1)$ and $(1, 1)$ with probability $p/3$. A *Pauli channel* on countably many qubits is a Radon probability measure on $\hat{V}$, the law of a random error applied to the qubits.

Proposition 5.4. *Let $\hat{V}$ be the group of labels of Pauli strings of arbitrary support, and for $0 < p \leq 1$ let ν_p be the depolarizing noise of rate p . Then the following hold:*

- (a) *the map $u \mapsto \omega(\cdot, u)$ is an isomorphism of topological groups from $\hat{V}$ onto V^* , which carries V onto the functionals of the actual errors;*
- (b) *the probes of the Pauli observable are the Pauli channels on countably many qubits, and the response of a mixture μ at a Pauli channel ν is the probability that a check drawn with law μ anticommutes with an independent error drawn with law ν ;*
- (c) *the depolarizing noise ν_p is concentrated on errors of infinite support, and for a label $a \in V$ that acts nontrivially on w qubits we have*

$$r_{\delta_a}(\nu_p) = \frac{1}{2} \left(1 - \left(1 - \frac{4p}{3} \right)^w \right).$$

Proof. (a) The map is a homomorphism. Write X_k and Z_k for the labels of the strings X and Z on the qubit k . Then the values $\omega(X_k, u)$ and $\omega(Z_k, u)$ are the two coordinates of u at the qubit k , so the map is injective, and every $\lambda \in V^*$ is the image of the vector u with these coordinates equal to $\lambda(X_k)$ and $\lambda(Z_k)$, because λ and $\omega(\cdot, u)$ are linear and agree on the basis $\{X_k, Z_k : k \in \mathbb{N}\}$ of V . The value $\omega(a, u)$ depends on finitely many coordinates of u , so the map is continuous, and a continuous bijection between compact Hausdorff spaces is a homeomorphism.

(b) By Lemma 5.1, Lemma 5.3 and (a), the probes are the Radon probability measures on $\hat{V}$, and $r_\mu(\nu) = \int \sum_a \mu(a) \omega(a, u) d\nu(u)$, which is the probability stated, by Fubini's theorem.

(c) Under ν_p , the events “the error acts nontrivially on the qubit k ” are independent and have probability p , so, by the lemma of Borel and Cantelli, almost every error acts nontrivially on infinitely many qubits. At each of the w qubits where P_a acts nontrivially, it anticommutes with two of the three nontrivial Pauli matrices, hence with the error with probability $2p/3$, independently of the other qubits. The strings anticommute exactly when an odd number of these w events occur, and this has probability $\frac{1}{2} \left(1 - \left(1 - \frac{4p}{3} \right)^w \right)$. $\square$

By Proposition 5.4, the noise of a memory on countably many qubits is a probe of the Pauli observable whose support consists of idealized errors: with probability one, the error acts on infinitely many qubits and is not an actual label. Alice's checks see such an error through finitely many qubits at a time, and the response of a check of weight w tends to $\frac{1}{2}$, the value of a fair coin, as w grows. In what follows we identify V^* with $\hat{V}$ and call its elements *idealized errors*. For $a \in V$, let χ_a be the character of V^* given by $\chi_a(\lambda) = (-1)^{\lambda(a)}$. Alice's pure responses are the characters of this compact group up to an affine change: we have $\lambda(a) = (1 - \chi_a(\lambda))/2$.

5.5. The Rado observable. The *Rado observable* is the diagonal observable of E on $\Gamma \times \Gamma$: Alice and Bob each hold a site of a random network of couplings, and the observable reports whether the two sites are coupled. By the extension property, it has the independence property already on pure preparations. We say that a diagonal observable has *freely independent* pure responses when $K_\varphi = \{0, 1\}^A$, that is, when every pattern of answers to all of Alice's labels is an idealized label of Bob.

Lemma 5.5. *For the Rado observable, the space K_E is the whole Cantor cube $\{0, 1\}^\Gamma$. Hence the Rado observable has freely independent pure responses, while the Pauli observable does not.*

Proof. A basic open subset of $\{0, 1\}^\Gamma$ prescribes adjacency to the vertices of a finite set Γ_1 and nonadjacency to those of a finite set Γ_0 , and the extension property gives a vertex outside $\Gamma_1 \cup \Gamma_0$ with that pattern. By Lemma 5.3, the space K_ω is the group V^* , which is a proper subset of $\{0, 1\}^V$. $\square$

The pure probes of the Rado observable are arbitrary sets of neighbors, that is, arbitrary subsets of Γ , and its probes are random sets of neighbors. The response of a mixture μ at a probe ν is the probability that a site drawn with law μ lies in an independent set of neighbors drawn with law ν ; it depends on ν only through the marginals $\nu\{\lambda : \lambda(a) = 1\}$, the probabilities that the individual sites a are neighbors. The Pauli observable sees more of a probe: its pure responses are, up to an affine change, the characters of the compact group of idealized errors, and products of characters are again characters. Sections 7 and 8 show what this difference does to the tree property.

6. THE THEORY OF MIXED PREPARATIONS HAS THE TREE PROPERTY OF THE SECOND KIND

The goal of this section is Theorem 1.1(C). The idea is a comparison of two averages. Let J be a finite set of labels and H a half of J . In the proof of Theorem 6.1 we show that the difference between the response to the uniform mixture on H and the response to the uniform mixture on $J \setminus H$ reaches its maximum 1 exactly when the probe answers 1 on H and 0 on $J \setminus H$, that two halves that meet in a quarter of J cannot both be nearly maximized, and that the independence property supplies a pure probe with any prescribed answers on finitely many disjoint blocks.

Theorem 6.1 (Mixed preparations and the tree property of the second kind).

Let C_φ be a diagonal observable, and assume that there are distinct labels $a_0, a_1, \dots$ in A such that for all finite disjoint $N_1, N_0 \subseteq \mathbb{N}$ there is a label $b \in B$ with $\varphi(a_n, b) = 1$ for $n \in N_1$ and $\varphi(a_n, b) = 0$ for $n \in N_0$. Then the affine formula

$$\psi(x; y_1, y_2) = R(y_1, x) - R(y_2, x),$$

with x of the sort of probes and y_1, y_2 of the sort of preparations, has the tree property of the second kind in the real-valued theory of $\mathcal{S}_\varphi$, with $k = 2$, $\alpha = 1/2$ and $\beta = 1$. In particular, a simple diagonal observable has no independence property.

Proof. We first build finite arrays in $\mathcal{S}_\varphi$ itself and then pass to an elementary extension by compactness.

Let $N \geq 1$ be given, and choose t with $2^t > N$. Partition $2^t N$ of the labels a_n into N blocks $J_0, \dots, J_{N-1}$ of size 2^t , and identify each block with $\mathbb{F}_2^t$. Choose distinct nonzero vectors $w_0, \dots, w_{N-1}$ in $\mathbb{F}_2^t$, and for $i, j < N$ let $H_{ij} \subseteq J_i$ be the set of $v \in \mathbb{F}_2^t$ with $w_j \cdot v = 1$.

Each H_{ij} has 2^{t-1} elements, and for $j \neq j'$ the set $H_{ij} \cap H_{ij'}$ has 2^{t-2} elements, since w_j and $w_{j'}$ are linearly independent. Let μ_{ij}^+ and μ_{ij}^- be the uniform probability measures on H_{ij} and on $J_i \setminus H_{ij}$.

For a probe ν , write $\pi_a(\nu) = \nu\{\lambda : \lambda(a) = 1\} \in [0, 1]$ for its marginal at the label a , so that $R(\mu, \nu) = \sum_a \mu(a)\pi_a(\nu)$ by Lemma 5.1. With $m = 2^t$, we have

$$\psi(\nu; \mu_{ij}^+, \mu_{ij}^-) = \frac{2}{m} \left(\sum_{a \in H_{ij}} \pi_a(\nu) - \sum_{a \in J_i \setminus H_{ij}} \pi_a(\nu) \right) \leq 1.$$

For $j \neq j'$ in the same row, adding the two values gives

$$\frac{4}{m} \left(\sum_{a \in H_{ij} \cap H_{ij'}} \pi_a(\nu) - \sum_{a \in J_i \setminus (H_{ij} \cup H_{ij'})} \pi_a(\nu) \right) \leq \frac{4}{m} \cdot \frac{m}{4} = 1,$$

so the two values are never both greater than $1/2$. For every $f: N \rightarrow N$, the hypothesis applied to the finite disjoint sets $\bigcup_i H_{if(i)}$ and $\bigcup_i (J_i \setminus H_{if(i)})$ gives a label b , and at the pure probe of $\varphi(\cdot, b)$ every value $\psi(\cdot; \mu_{if(i)}^+, \mu_{if(i)}^-)$ with $i < N$ equals 1.

Thus, for every N , the structure $\mathcal{S}_\varphi$ satisfies the following conditions on $2N^2$ preparations $y_{ij}^\pm$: for every $f: N \rightarrow N$, the value of $\sup_x \min_{i < N} \psi(x; y_{if(i)}^+, y_{if(i)}^-)$ is 1; and for all i and $j \neq j'$, the value of $\sup_x \min(\psi(x; y_{ij}^+, y_{ij}^-), \psi(x; y_{ij'}^+, y_{ij'}^-))$ is at most $1/2$. These are finitely many closed conditions for each N , and the conditions for N are inherited by subarrays. By compactness, an $\aleph_1$ -saturated elementary extension of $\mathcal{S}_\varphi$ contains an array $(y_{ij}^\pm)_{i,j \in \mathbb{N}}$ satisfying them for all N ; by saturation, the suprema along paths are attained. This is the tree property of the second kind, which is incompatible with simplicity in real-valued logic [BY03].

It remains to be shown that a simple diagonal observable has no independence property. The hypothesis of the theorem is the independence property of the pure responses, in the form used in the second paper: actual labels are dense in K_φ , and a finite pattern of answers to the labels a_n prescribes a clopen subset of K_φ , so the hypothesis says exactly that the functions $\lambda \mapsto \lambda(a_n)$ form an independent sequence on K_φ . For a diagonal observable, the pure and the mixed responses have the independence property together: a mixed response is an average of pure ones, and if the pure responses have no independent sequence, then, by the dichotomy of Rosenthal, every sequence of them has a subsequence that converges pointwise on K_φ , hence is weakly Cauchy in $C(K_\varphi)$ by dominated convergence, and by the theorem of Stegall, as in the second paper, so does every sequence in the closed absolutely convex hull of the pure responses, which contains the mixed responses. Hence, if a diagonal observable has the independence property, the hypothesis holds, and the real-valued theory of $\mathcal{S}_\varphi$ has the tree property of the second kind. $\square$

In Sections 7 and 8 we determine, for each paradigm, whether actual preparations already witness the tree property of the second kind: for the Pauli observable they do, and for the Rado observable they do not.

7. THE PAULI OBSERVABLE: SYNDROME SECTORS AND THE FIRST CHECK TO FIRE

The goal of this section is Theorem 1.1(D): for the Pauli observable, actual preparations witness the tree property of the second kind. The tool is the formula for the projector onto a syndrome sector as an average over a group of checks, which we first state for responses. The idea of the theorem is then to let a single mixture ask a question about many checks at once. Let a block of independent checks $g_0, g_1, \dots$ be given, and let Alice draw a Pauli

string uniformly from the coset $g_j + \text{span}(g_0, \dots, g_{j-1})$. In the proof of Theorem 7.2 we show that if the checks before g_j are silent on Bob's error and g_j fires, every string in the coset anticommutes with the error, and that if one of the earlier checks fires, exactly half of them do. Two different values of j cannot both describe the first check to fire, and blocks of disjoint checks are independent (see Lemma 3.2).

7.1. Syndrome sectors. Let $g_1, \dots, g_d \in V$ be linearly independent checks, with $\text{span } G$, and let a syndrome $\sigma \in \mathbb{F}_2^d$ be given. We regard σ as the linear functional on G with $\sigma(g_i) = \sigma_i$, and we call the set $\Sigma_\sigma = \{\lambda \in V^* : \lambda(g_i) = \sigma_i \text{ for } i \leq d\}$ of idealized errors with syndrome σ its *syndrome sector*. For $s \in G$ we write r_s for the pure response r_{δ_s} .

Proposition 7.1. *The indicator of the syndrome sector Σ_σ satisfies, at every pure probe $\lambda \in V^*$,*

$$\mathbf{1}_{\Sigma_\sigma}(\lambda) = 2^{-d} \sum_{s \in G} (-1)^{\sigma(s)} \chi_s(\lambda) = 2^{-d} \sum_{s \in G} (-1)^{\sigma(s)} - 2^{1-d} \sum_{s \in G} (-1)^{\sigma(s)} r_s(\lambda).$$

Consequently, if $\sigma \neq 0$, then $\mathbf{1}_{\Sigma_\sigma} = r_{\mu_\sigma^1} - r_{\mu_\sigma^0}$ on pure probes, where μ_σ^1 and μ_σ^0 are the uniform probability measures on $\{s \in G : \sigma(s) = 1\}$ and on $\{s \in G : \sigma(s) = 0\}$; and $\mathbf{1}_{\Sigma_0} = 1 - 2r_{\mu_G}$, where μ_G is the uniform probability measure on G . In particular, for every probe ν , the probability that an error drawn from the Pauli channel ν has syndrome σ is $r_{\mu_\sigma^1}(\nu) - r_{\mu_\sigma^0}(\nu)$ if $\sigma \neq 0$, and $1 - 2r_{\mu_G}(\nu)$ if $\sigma = 0$.

Proof. The map $s \mapsto (-1)^{\sigma(s)} \chi_s(\lambda) = (-1)^{(\sigma+\lambda)(s)}$ is a character of the finite group G ; its sum over G is 2^d if the linear functional $\sigma + \lambda$ vanishes on G , that is, if $\lambda \in \Sigma_\sigma$, and 0 otherwise, since a nonzero functional on G takes the value 1 on exactly half of G . This is the first equality, and the second follows from $\chi_s = 1 - 2r_s$. If $\sigma \neq 0$, the first sum on the right vanishes, and the two level sets of σ have 2^{d-1} elements each; grouping the second sum by the value of σ gives $r_{\mu_\sigma^1} - r_{\mu_\sigma^0}$. If $\sigma = 0$, the first sum is 2^d , and the second is $2^d r_{\mu_G}$. Both sides of each identity are affine in the probe, so integrating against ν gives the last statement, by Proposition 5.4(b). $\square$

When the checks commute pairwise and carry signs, the operator $2^{-d} \sum_{s \in G} (-1)^{\sigma(s)} P_s$ is the projector onto the syndrome sector σ of the code [Got97], and Proposition 7.1 gives the probability of the outcome σ when the checks are measured on a code state corrupted by the channel ν . The proposition requires no commutation: the checks need only be linearly independent. The coefficients of the indicator, as an affine combination of Alice's responses, have total absolute value at most $1 + 2 = 3$, whatever the number d of checks.

7.2. The first check to fire. The next theorem arranges the syndrome sectors of nested blocks of checks in an array that witnesses the tree property of the second kind.

Theorem 7.2 (The first check to fire). *Let $(g_{im})_{i,m \in \mathbb{N}}$ be linearly independent vectors of V , put $W_{ij} = \text{span}(g_{im} : m < j)$, and let μ_{ij} be the uniform probability measure on the coset $g_{ij} + W_{ij}$. For the Pauli observable, the following hold:*

- (a) *for every $f : \mathbb{N} \rightarrow \mathbb{N}$ and every n , there is an error $e \in V$ with $r_{\mu_{if(i)}}(e) = 1$ for all $i < n$, and there is an idealized error $\lambda \in V^*$ with $r_{\mu_{if(i)}}(\lambda) = 1$ for all i ;*
- (b) *for every i , every $j \neq j'$ and every probe $\nu \in \mathcal{P}(V^*)$, we have $\min(r_{\mu_{ij}}(\nu), r_{\mu_{ij'}}(\nu)) \leq 2/3$, and the bound $2/3$ is attained.*

Consequently, the responses of Alice have the tree property of the second kind, with $k = 2$, $\alpha = 2/3$ and $\beta = 1$.

Proof. We first compute the response of μ_{ij} at a pure probe $\lambda \in V^*$. If λ does not vanish on W_{ij} , then the functional λ takes the value 1 on exactly half of the coset $g_{ij} + W_{ij}$, since a nonzero functional on a finite $\mathbb{F}_2$ -space is balanced, and translation by g_{ij} either preserves or exchanges its level sets. If λ vanishes on W_{ij} , then the functional λ is constant on the coset, with value $\lambda(g_{ij})$. Hence

$$(1) \quad r_{\mu_{ij}}(\lambda) = \begin{cases} 1 & \text{if } \lambda(g_{im}) = 0 \text{ for } m < j \text{ and } \lambda(g_{ij}) = 1, \\ 1/2 & \text{if } \lambda(g_{im}) = 1 \text{ for some } m < j, \\ 0 & \text{otherwise.} \end{cases}$$

In words, the response is 1 exactly when g_{ij} is the first check of block i to fire.

(a) Let f and n be given. The conditions $\lambda(g_{im}) = 0$ for $m < f(i)$ and $\lambda(g_{if(i)}) = 1$, for $i < n$, concern finitely many linearly independent vectors, so Lemma 3.2 gives an error e with $\omega(\cdot, e)$ satisfying them, and (1) gives $r_{\mu_{if(i)}}(e) = 1$ for all $i < n$. The conditions for all i together define a nonempty closed subset of V^* , since the vectors g_{im} are linearly independent; any λ in it works.

(b) By symmetry, we may assume that $j < j'$. By (1), the pair $(r_{\mu_{ij}}(\lambda), r_{\mu_{ij'}}(\lambda))$ at a pure probe takes one of the values

$$(1, \tfrac{1}{2}), \quad (\tfrac{1}{2}, \tfrac{1}{2}), \quad (0, \tfrac{1}{2}), \quad (0, 1), \quad (0, 0).$$

Indeed, if the first check of block i to fire among $g_{i0}, \dots, g_{ij'}$ has index m , the pair is $(\frac{1}{2}, \frac{1}{2})$ for $m < j$, $(1, \frac{1}{2})$ for $m = j$, $(0, \frac{1}{2})$ for $j < m < j'$ and $(0, 1)$ for $m = j'$; if none fires, the pair is $(0, 0)$. Responses are affine in the probe, so at a mixed probe ν the pair lies in the convex hull of these five points. On that hull the minimum of the two coordinates is at most $2/3$, with equality only at $\frac{2}{3}(1, \frac{1}{2}) + \frac{1}{3}(0, 1)$, which is the pair of the probe $\frac{2}{3}\delta_\lambda + \frac{1}{3}\delta_{\lambda'}$ with g_{ij} firing first under λ and $g_{ij'}$ under λ' .

Statements (a) and (b) are the two clauses of Definition 5.2 for the array $(r_{\mu_{ij}})$, with $k = 2$, $\alpha = 2/3$ and $\beta = 1$. $\square$

Remark 7.3. For qudits of prime dimension p , the labels form a symplectic space over the field $\mathbb{F}_p$ with p elements, and the response of a label at an error is 1 when their commutation phase is nontrivial and 0 otherwise. The proof of Theorem 7.2 adapts as follows. A nonzero functional on a finite $\mathbb{F}_p$ -space takes each value equally often, so in (1) the response $\frac{1}{2}$ becomes $q = \frac{p-1}{p}$. For $j < j'$, the pair of responses at a pure probe takes one of the values $(1, q)$, (q, q) , $(0, q)$, $(0, 1)$ and $(0, 0)$. The point (q, q) lies below the segment from $(1, q)$ to $(0, 1)$, since $q < 1 - q + q^2$, and on that segment the two coordinates are equal at the point with first coordinate $\frac{p}{p+1}$. Hence the bound $\frac{2}{3}$ of Theorem 7.2(b) becomes $\frac{p}{p+1}$, and it is attained.

In the language of Proposition 7.1, the event “the check g_{ij} fires first in block i ” is the syndrome sector of the checks $g_{i0}, \dots, g_{ij}$ with syndrome $(0, \dots, 0, 1)$, whose two level sets are the coset $g_{ij} + W_{ij}$ and the group W_{ij} . Hence its indicator is $r_{\mu_{ij}} - r_{\mu_{ij}^0}$ on pure probes, where μ_{ij}^0 is the uniform probability measure on W_{ij} . Theorem 7.2 uses only linear independence; when the checks commute pairwise, each W_{ij} is the set of labels of a stabilizer group, and μ_{ij} draws uniformly from a coset of it. The simplest instance is the repetition code.

Example 7.4. Index the qubits by pairs $(i, m) \in \mathbb{N}^2$, so that each block i is a chain of qubits, and let $g_{im} = Z_{i,m}Z_{i,m+1}$ be the check of the bit-flip repetition code on the neighboring qubits m and $m+1$ of the chain i . These checks commute pairwise and are linearly independent. For a bit flip $e = \prod X_{i,m}^{x_{i,m}}$, we have $\omega(g_{im}, e) = x_{i,m} + x_{i,m+1}$, which is 1 exactly when e has a domain wall between the qubits m and $m+1$ of the chain i . By (1), the response of μ_{ij} at e is 1 exactly when the first domain wall of e on the chain i lies between the qubits j and $j+1$. Theorem 7.2 then reads as follows: the chains are independent code blocks of a memory, an error cannot have two first domain walls on the same chain, and the positions of the first domain walls on finitely many chains can be prescribed independently by a single error. The bit flips on the chain i with zero syndrome are the identity and the flip X of the whole chain, which acts on infinitely many qubits: the logical bit flip of an infinite chain is an idealized error, an element of $\hat{V} \setminus V$.

Remark 7.5. Let us notice that the mechanism behind Theorem 7.2 is Fourier analysis on the compact group V^* : since the characters of V^* separate points and are closed under multiplication, the linear span of Alice’s responses and the constants is dense in $C(V^*)$ by the Stone–Weierstrass theorem; it is an algebra, since $2r_a r_b = r_a + r_b - r_{a+b}$ for the pure responses r_a . By Proposition 7.1, the indicators of syndrome sectors, which are functions of supremum norm one, lie in this span with coefficients of total absolute value at most 3. On the pure labels, by contrast, a formula of (V, ω) speaks about a bounded number of checks, and Theorem 3.3 shows that no formula has the tree property.

Theorem 6.1 shows that the real-valued theory of $\mathcal{S}_\omega$ is not simple, although the classical theory of (V, ω) is; Theorem 7.2 places the tree in $\mathcal{S}_\omega$ itself, with actual preparations as parameters and the atomic formula as the formula.

8. THE RADO OBSERVABLE: MARGINALS AND IDEALIZED PREPARATIONS

The goal of this section is Theorem 1.1(E). By Theorem 6.1, the Rado observable has the tree property of the second kind, as the Pauli observable does. In Theorem 8.1 and Corollary 8.2 we show that its tree lives among idealized preparations, because conditions on marginals given by actual mixtures are compatible on disjoint blocks. After Corollary 8.2 we note that the same holds for the random bipartite graph, the Fraïssé limit of the finite bipartite graphs with the two sides as two sorts, which has exactly the shape of a bipartite observable, with Alice’s labels and Bob’s labels as the two sorts.

8.1. Affine combinations of responses. For an observable with freely independent pure responses, the response of a mixture depends on a probe only through its marginals $\pi_a(\nu) = \nu\{\lambda : \lambda(a) = 1\}$, as for the Rado observable in Section 5.5, and every vector of marginals in $[0, 1]^A$ is attained by a product measure. A bounded affine combination of responses is therefore a function

$$h_{c,\gamma}(\nu) = \gamma + \sum_{a \in A} c(a) \pi_a(\nu), \quad c \in \ell^1(A),$$

and its supremum over the probe space is $\gamma + \sum_a c(a)^+$, attained at the pure probe $\lambda = \mathbf{1}_{\{c>0\}}$. Two such functions can only conflict through the labels on which both put weight. In the proof of Theorem 8.1 below we show that an infinite sequence of them must, after a gliding hump, put its weight on disjoint blocks of labels, apart from a common part, so that a single pure probe nearly maximizes all of them.

Theorem 8.1 (Compatibility of affine conditions). *Let C_φ be a diagonal observable with freely independent pure responses, and let $\kappa > 0$ be given. Let $\mathcal{R}_\kappa$ be the family of functions $\gamma + \sum_{i \leq n} c_i r_{\mu_i}$ on $\mathcal{P}(\{0, 1\}^A)$, where $n \in \mathbb{N}$, the μ_i are mixtures of Alice and $|\gamma| + \sum_i |c_i| \leq \kappa$. Then the family $\mathcal{R}_\kappa$ contains no uniformly k -inconsistent sequence, for any k . In particular, no tree property is witnessed in $\mathcal{R}_\kappa$.*

Proof. Let (h_j) be a sequence in $\mathcal{R}_\kappa$ and let $\alpha < \beta$ be such that each h_j takes a value $\geq \beta$. We show that for every k there are $j_1 < \dots < j_k$ and a pure probe λ with $h_{j_\ell}(\delta_\lambda) > \alpha$ for all ℓ .

Write $h_j = h_{c_j, \gamma_j}$ with $c_j = \sum_i c_{j,i} \mu_{j,i}$, so that $\|c_j\|_1 \leq \kappa$. Since A is countable, a diagonal argument gives a subsequence, still denoted (c_j) , that converges pointwise on A to some c ; by Fatou's lemma, the limit satisfies $\|c\|_1 \leq \kappa$. Put $\eta = (\beta - \alpha)/5$ and choose a finite set $F \subseteq A$ with $\|c \upharpoonright (A \setminus F)\|_1 < \eta$. Put $F_0 = F$, and choose inductively indices $j_1 < \dots < j_k$ and finite sets $F_0 \subseteq F_1 \subseteq \dots \subseteq F_k$ such that

$$\|(c_{j_\ell} - c) \upharpoonright F_{\ell-1}\|_1 < \eta \quad \text{and} \quad \|c_{j_\ell} \upharpoonright (A \setminus F_\ell)\|_1 < \eta;$$

the first condition holds for all large indices by pointwise convergence, and the second for a large enough finite set. Define $\lambda \in \{0, 1\}^A$ by $\lambda(a) = 1$ exactly when $a \in F$ and $c(a) > 0$, or $a \in F_\ell \setminus F_{\ell-1}$ and $c_{j_\ell}(a) > 0$ for some $\ell \leq k$.

Let $\ell \leq k$ be given. The deficit of λ for h_{j_ℓ} is

$$\sup h_{j_\ell} - h_{j_\ell}(\delta_\lambda) = \sum_{\lambda(a)=0} c_{j_\ell}(a)^+ + \sum_{\lambda(a)=1} c_{j_\ell}(a)^-,$$

and we bound its contribution on four disjoint pieces of A . On $F_\ell \setminus F_{\ell-1}$ the sign pattern of λ is that of c_{j_ℓ} , and the contribution is 0. On F it is that of c , and a label where the signs of c_{j_ℓ} and c disagree contributes at most $|c_{j_\ell}(a) - c(a)|$; the total is less than η . On $F_{\ell-1} \setminus F$ the contribution is at most $\|c_{j_\ell} \upharpoonright (F_{\ell-1} \setminus F)\|_1 \leq \|(c_{j_\ell} - c) \upharpoonright F_{\ell-1}\|_1 + \|c \upharpoonright (A \setminus F)\|_1 < 2\eta$. On $A \setminus F_\ell$ it is less than η . Hence $h_{j_\ell}(\delta_\lambda) > \sup h_{j_\ell} - 4\eta \geq \beta - 4\eta > \alpha$. $\square$

Corollary 8.2. *For the Rado observable, no tree property of a bounded affine combination of responses is witnessed by actual preparations, although the real-valued theory of $\mathcal{S}_E$ has the tree property of the second kind.*

Proof. By Lemma 5.5, the Rado observable has freely independent pure responses, and Theorem 8.1 applies. The second statement is Theorem 6.1. $\square$

The finite arrays in the proof of Theorem 6.1 are made of actual preparations, and their rows have any prescribed finite length; by Theorem 8.1, the rows cannot be made infinite in $\mathcal{S}_E$. In an $\aleph_1$ -saturated elementary extension, the rows are infinite, by compactness, and Theorem 8.1 shows that the entries of an infinite row cannot all be actual mixtures: the arrays that witness the tree property of the second kind in Theorem 6.1 need idealized preparations. The same holds for the random bipartite graph, whose observable has freely independent pure responses by its extension property.

8.2. Quantifier-free formulas. Theorem 8.1 concerns affine combinations of responses. In the real-valued logic of the second paper, a quantifier-free formula in a variable x of probe, built from the response with parameters of preparation, is a continuous function of finitely many responses, since the responses take values in $[0, 1]$ and the connectives are continuous functions. The next theorem is the analogue of Theorem 8.1 for these formulas. Its proof uses the same gliding hump, with one change: a continuous function of responses is no longer maximized by reading off signs (we show this after the proof), so instead of building the probe from signs we glue maximizers of the individual instances on disjoint blocks of labels.

Theorem 8.3 (Compatibility of quantifier-free conditions). *Let C_φ be a diagonal observable with freely independent pure responses, let $n \geq 1$ be an integer, and let $u: [0, 1]^n \rightarrow \mathbb{R}$ be a continuous function. Let $\mathcal{R}_u$ be the family of functions*

$$h(\nu) = u(r_{\mu_1}(\nu), \dots, r_{\mu_n}(\nu)) \quad (\nu \in \mathcal{P}(\{0, 1\}^A)),$$

where $\mu_1, \dots, \mu_n$ range over the mixtures of Alice. Then the family $\mathcal{R}_u$ contains no uniformly k -inconsistent sequence, for any k . In particular, no tree property of the quantifier-free formula $u(R(y_1, x), \dots, R(y_n, x))$ in the probe variable x is witnessed by actual preparations.

Proof. For $\nu \in \mathcal{P}(\{0, 1\}^A)$ write $\pi(\nu) \in [0, 1]^A$ for its vector of marginals, as in the proof of Theorem 8.1. Every $h \in \mathcal{R}_u$ factors as $h(\nu) = H(\pi(\nu))$ with

$$H(q) = u(\langle \mu_1, q \rangle, \dots, \langle \mu_n, q \rangle), \quad \langle \mu, q \rangle = \sum_{a \in A} \mu(a) q_a,$$

and every $q \in [0, 1]^A$ is the vector of marginals of a product measure. It therefore suffices to work on the cube $[0, 1]^A$.

Let (h_j) be a sequence in $\mathcal{R}_u$, with parameters $\mu_{j,1}, \dots, \mu_{j,n}$, and let $\alpha < \beta$ be such that each h_j takes a value $\geq \beta$. We show that for every k there are $j_1 < \dots < j_k$ and $q \in [0, 1]^A$ with $H_{j_\ell}(q) > \alpha$ for all ℓ . Since u is uniformly continuous, there is $\theta > 0$ such that $|u(s) - u(t)| < \beta - \alpha$ whenever $\max_i |s_i - t_i| < \theta$. For each j choose $q^{(j)} \in [0, 1]^A$ with $H_j(q^{(j)}) \geq \beta$. Since A is countable and the cube is compact and metrizable, a diagonal argument gives a subsequence, still indexed by j , along which $\mu_{j,i} \rightarrow \mu_i$ pointwise on A for each $i \leq n$ and $q^{(j)} \rightarrow q^\infty$ in $[0, 1]^A$. By Fatou's lemma, each limit μ_i has total mass at most 1. Put $\eta = \theta/5$ and choose a finite set $F \subseteq A$ with $\mu_i(A \setminus F) < \eta$ for all $i \leq n$.

Put $F_0 = F$, and choose inductively indices $j_1 < \dots < j_k$ and finite sets $F_0 \subseteq F_1 \subseteq \dots \subseteq F_k$ such that, for all $i \leq n$,

$$\sum_{a \in F_{\ell-1}} |\mu_{j_\ell, i}(a) - \mu_i(a)| < \eta, \quad \sum_{a \in F} |q_a^{(j_\ell)} - q_a^\infty| < \eta, \quad \mu_{j_\ell, i}(A \setminus F_\ell) < \eta.$$

The first two conditions hold for all large indices, by convergence on the finite sets $F_{\ell-1}$ and F , and the third for a large enough finite set F_ℓ . Define $q \in [0, 1]^A$ by $q_a = q_a^\infty$ for $a \in F$, $q_a = q_a^{(j_\ell)}$ for $a \in F_\ell \setminus F_{\ell-1}$ and $\ell \leq k$, and $q_a = 0$ for $a \notin F_k$.

Let $\ell \leq k$ and $i \leq n$ be given. We compare $\langle \mu_{j_\ell, i}, q \rangle$ with $\langle \mu_{j_\ell, i}, q^{(j_\ell)} \rangle$, on the same four pieces of A as in the proof of Theorem 8.1; on each piece the two vectors differ by at most 1 in each coordinate. On $F_\ell \setminus F_{\ell-1}$ they agree. On F the difference is at most $\sum_{a \in F} |q_a^\infty - q_a^{(j_\ell)}| < \eta$, since $\mu_{j_\ell, i}(a) \leq 1$. On $F_{\ell-1} \setminus F$ it is at most $\mu_{j_\ell, i}(F_{\ell-1} \setminus F) \leq \sum_{a \in F_{\ell-1}} |\mu_{j_\ell, i}(a) - \mu_i(a)| + \mu_i(A \setminus F)$

$F) < 2\eta$. On $A \setminus F_\ell$ it is at most $\mu_{j_\ell, i}(A \setminus F_\ell) < \eta$. Hence $|\langle \mu_{j_\ell, i}, q \rangle - \langle \mu_{j_\ell, i}, q^{(j_\ell)} \rangle| < 4\eta < \theta$ for every $i \leq n$, and the choice of θ gives $H_{j_\ell}(q) > H_{j_\ell}(q^{(j_\ell)}) - (\beta - \alpha) \geq \alpha$. $\square$

The proof of Theorem 8.1 gives slightly more for affine combinations: there the common witness can be taken pure. For a general connective u the witness is the product probe whose vector of marginals is q , and a mixed witness is needed in general: for $n = 1$, $\mu = \delta_a$ and $u(t) = 1 - |2t - 1|$, the instance vanishes at every pure probe and equals 1 at the probes with $q_a = 1/2$.

Let us notice that the difference between the two paradigms lies entirely in the probe space: on $\{0, 1\}^A$, affine combinations of coordinates see only marginals, and conditions on marginals of disjoint blocks are compatible. On the group V^* , the linear relations among the labels make products of coordinates affine combinations of coordinates, and the indicators of syndrome sectors become affine combinations of responses (see Proposition 7.1).

9. AFFINE LOGIC AND CORRELATED PREPARATIONS

The goal of this section is to place the results in affine logic, in the form developed by Ben Yaacov, Ibarlucía and Tsankov [BYIT24], whose connectives are the affine maps and in which a mixture of structures is a direct integral. Formula values in a direct integral are integrals of formula values, which is the rule by which a response is computed on a mixed preparation. In this logic, a structure is *extremal* when it realizes only extreme types. Ben Yaacov, Ibarlucía and Tsankov [BYIT24] showed that every model of an affine theory is an affine elementary substructure of a direct integral of extremal models, and that every model of a simplicial theory, one whose type spaces are Choquet simplices, is such a direct integral, uniquely.

Ben Yaacov and Ibarlucía [BYI26] proved that an affine formula stable in almost every member of a measurable field of structures is stable in its direct integral, and that an affine theory is stable exactly when its extremal models are stable. Their definition of stability is the double limit condition used in the first paper of this series [DI26a]. The mixing of this paper behaves differently with respect to simplicity: the pure structures of the Pauli group and of the Rado graph are simple, and, by Theorem 6.1, the structures of their observables have the tree property of the second kind.

The direct integral allows the elements in the two variables of a formula to depend on the same random parameter. For observables this means that Alice and Bob share randomness: they prepare a separable state, a mixture of product states, rather than a single product state. The randomizations of Ben Yaacov [BY13] are of this kind, while Theorem 6.1 concerns product preparations.

10. QUESTIONS

Question 10.1. Which diagonal observables with the independence property have the tree property of the second kind witnessed by actual preparations? Theorems 7.2 and 8.1 and Proposition 7.1 suggest a criterion in terms of the linear span of the responses in $C(K_\varphi)$: whether it contains, with coefficients of bounded total absolute value, the indicators of an infinite independent sequence of nested clopen sets.

Question 10.2. For qudits of prime dimension p , the labels form a symplectic space over the field with p elements and the observable reports a nonzero commutation phase. By Remark 7.3, the proof of Theorem 7.2 adapts to qudits. For continuous variables, the commutation phase of a translation by x and a momentum kick by k is the interference fringe

e^{ikx} , which is not NIP by the second paper. Where does the unbounded fringe lie with respect to simplicity?

Question 10.3. In classical model theory, a simple theory without the independence property is stable. Is a simple observable stable? By Theorem 6.1, a simple diagonal observable has no independence property, so the question is whether the unstable observables of the second paper, such as the comparison of photon numbers, have a tree property once preparations are mixed.

Question 10.4. Does Theorem 6.1 hold for every observable whose responses have the independence property, not only for diagonal ones? The proof uses the values 0 and 1 of pure responses; for general observables they are replaced by two levels $a < b$, and the averages must be controlled with a gap.

Question 10.5. Is there a theorem of Kim and Pillay for affine logic, characterizing simple affine theories by an independence relation on extremal types?

HISTORICAL REMARKS

Simple theories were introduced by Shelah [She80] in 1980, together with the tree properties of the first and second kind; the motivating example was T_{ind} , the theory of the countable random graph (see [She80, GIL02]). Kim [Kim98] proved in 1998 that forking is symmetric in simple theories, and Kim and Pillay [KP97] proved the independence theorem and the characterization of simplicity by an independence relation in 1997. Ben Yaacov [BY03] extended the theory to compact abstract theories and to continuous logic. The Rado graph was described by Ackermann [Ack37] in 1937 and by Rado [Rad64] in 1964; by a theorem of Erdős and Rényi [ER63] from 1963, it is the almost sure isomorphism type of the random graph on countably many vertices. Fraïssé [Fra54] introduced his limits in 1954, and the theorem of Witt on isometries of forms is in Artin's *Geometric Algebra* [Art57]. Infinite-dimensional vector spaces over finite fields with a nondegenerate form are among the smoothly approximable structures of Cherlin and Hrushovski [CH03]. Granger [Gra99] studied vector spaces with a generic bilinear form over an arbitrary field in his thesis of 1999 and showed that, over an infinite field carried as a second sort, the theory is not simple; Chernikov and Ramsey [CR16] proved in 2016 that over an algebraically closed field it is NSOP₁. The symplectic description of the Pauli group and of stabilizer codes is due to Gottesman [Got97] and to Calderbank, Rains, Shor and Sloane [CRSS98]; the expression of the projector onto a syndrome sector as an average over the stabilizer group, behind Proposition 7.1, is standard in the stabilizer formalism of Gottesman. Keisler [Kei99] introduced randomizations in 1999; Ben Yaacov and Keisler [BYK09] treated them as metric structures, and Ben Yaacov [BY13] proved in 2013 that the randomization of a theory with the independence property has TP₂. Direct integrals and extremal models in affine logic are due to Ben Yaacov, Ibarlucía and Tsankov [BYIT24], and stability in affine logic to Ben Yaacov and Ibarlucía [BYI26]. The first two papers of this series are [DI26a, DI26b].

DECLARATIONS

Competing interests. The authors have no competing interests to declare.

Data availability. No data were generated or analyzed in this work.

REFERENCES

- [Ack37] W. Ackermann, *Die Widerspruchsfreiheit der allgemeinen Mengenlehre*, Math. Ann. **114** (1937), 305–315.

- [Art57] E. Artin, *Geometric Algebra*, Interscience, New York, 1957.
- [BDSW96] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed-state entanglement and quantum error correction*, Phys. Rev. A **54** (1996), 3824–3851.
- [BY03] I. Ben Yaacov, *Simplicity in compact abstract theories*, J. Math. Log. **3** (2003), 163–191.
- [BY13] Itai Ben Yaacov, *On theories of random variables*, Isr. J. Math. **194** (2013), 957–1012.
- [BYI26] I. Ben Yaacov and T. Ibarlucía, *Stability in affine logic*, Model Theory **5** (2026), 1–34.
- [BYIT24] I. Ben Yaacov, T. Ibarlucía, and T. Tsankov, *Extremal models and direct integrals in affine logic*, arXiv:2407.13344, 2024.
- [BYK09] Itai Ben Yaacov and H. Jerome Keisler, *Randomizations of models as metric structures*, Confluentes Math. **1** (2009), no. 2, 197–223. MR 2561997
- [CH03] G. Cherlin and E. Hrushovski, *Finite Structures with Few Types*, Annals of Mathematics Studies, vol. 152, Princeton University Press, 2003.
- [CR16] A. Chernikov and N. Ramsey, *On model-theoretic tree properties*, J. Math. Log. **16** (2016), 1650009.
- [CRSS98] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, *Quantum error correction via codes over $GF(4)$* , IEEE Trans. Inform. Theory **44** (1998), 1369–1387.
- [DDM03] J. Dehaene and B. De Moor, *Clifford group, stabilizer states, and linear and quadratic operations over $GF(2)$* , Phys. Rev. A **68** (2003), 042318.
- [DI26a] E. Dueñez and J. Iovino, *Classification of bipartite observables I: stable and uniformly stable observables*, preprint; arXiv:2610.11002, 2026.
- [DI26b] ———, *Classification of bipartite observables II: NIP observables and the Todorčević trichotomy*, preprint; arXiv:2610.12908, 2026.
- [ER63] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta Math. Acad. Sci. Hungar. **14** (1963), 295–315.
- [Fra54] R. Fraïssé, *Sur l’extension aux relations de quelques propriétés des ordres*, Ann. Sci. École Norm. Sup. **71** (1954), 363–388.
- [GIL02] Rami Grossberg, José N. Iovino, and Olivier Lessmann, *A primer of simple theories*, Arch. Math. Logic **41** (2002), no. 6, 541–580. MR MR1923196 (2003f:03043)
- [Got97] D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. thesis, California Institute of Technology, 1997, arXiv:quant-ph/9705052.
- [Gra99] N. Granger, *Stability, simplicity and the model theory of bilinear forms*, Ph.D. thesis, University of Manchester, 1999.
- [Kei99] H. Jerome Keisler, *Randomizing a model*, Adv. Math. **143** (1999), no. 1, 124–158. MR 1680650
- [Kim98] B. Kim, *Forking in simple unstable theories*, J. London Math. Soc. (2) **57** (1998), 257–267.
- [KP97] Byunghan Kim and Anand Pillay, *Simple theories*, Ann. Pure Appl. Logic **88** (1997), no. 2-3, 149–164, Joint AILA-KGS Model Theory Meeting (Florence, 1995). MR MR1600895 (99b:03049)
- [LMPZ96] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, *Perfect quantum error correcting code*, Phys. Rev. Lett. **77** (1996), 198–201.
- [Rad64] R. Rado, *Universal graphs and universal functions*, Acta Arith. **9** (1964), 331–340.
- [She80] S. Shelah, *Simple unstable theories*, Annals of Math. Logic **19** (1980), 177–203.
- [She90] S. Shelah, *Classification theory and the number of nonisomorphic models*, second ed., Studies in Logic and the Foundations of Mathematics, vol. 92, North-Holland Publishing Co., Amsterdam, 1990. MR MR1083551 (91k:03085)

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT SAN ANTONIO, ONE UTSA CIRCLE,
SAN ANTONIO, TX 78249, USA

Email address: Eduardo.Duenes@utsa.edu

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT SAN ANTONIO, ONE UTSA CIRCLE,
SAN ANTONIO, TX 78249, USA

Email address: jose.iovino@utsa.edu