Shelah's dividing lines in quantum measurement

30 variations in three parts

José Iovino

Department of Mathematics
The University of Texas at San Antonio

2020 *Mathematics Subject Classification.* Primary 03C45, 81P15; Secondary 03C66, 46L30, 26A21, 81P73.

Author's address: Department of Mathematics, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA; jose.iovino@utsa.edu.

This version: November 16, 2026. The most recent version of this memoir is available at https://utsa-quantum-math.github.io/resources/dividing-lines-quantum-measurement/; it may supersede the version deposited on arXiv.

Copyright © 2026 José Iovino. All rights reserved.

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author, except for brief quotations in scholarly work with full attribution.

This work is protected by the copyright laws of the United States and by the Berne Convention for the Protection of Literary and Artistic Works.

In homage to Saharon Shelah and Stevo Todorčević

Contents

Abstract vii

Introduction 1

Aria 6

Variatio 1. a 1 Clav. · *Observables as structures* 9

Part I. Stability 15

Variatio 2. a 1 Clav. · *The order property* 16

Variatio 3. a 1 Clav. Canone all'Unisuono · *Formulas and observables* 20

Variatio 4. a 1 Clav. · *Hilbert space* 23

Variatio 5. a 1 ô vero 2 Clav. · *Stability in seven languages* 26

Variatio 6. a 1 Clav. Canone alla Seconda · *The independent product* 30

Variatio 7. a 1 ô vero 2 Clav. Al tempo di Giga · *Discrimination and learning* 34

Variatio 8. a 2 Clav. · *Bounds independent of the dimension* 37

Variatio 9. a 1 Clav. Canone alla Terza · *Mixtures simulate idealized preparations* 39

Variatio 10. a 1 Clav. Fughetta · *Countable randomization does not detect instability* 43

Part II. NIP 47

Variatio 11. a 2 Clav. · *The independence property* 48

Variatio 12. Canone alla Quarta · *The rational order* 51

Variatio 13. a 2 Clav. · *The dictionary of NIP* 53

Variatio 14. a 2 Clav. · *Todorčević's trichotomy* 56

Variatio 15. a 1 Clav. Canone alla Quinta. Andante · *The heptachotomy* 60

Variatio 16. a 1 Clav. Ouverture · *Physical orders in every class* 65

Variatio 17. a 2 Clav. · *Orders of limits in phase transitions* 68

Variatio 18. a 1 Clav. Canone alla Sesta · *A quantum ferromagnet in the classification* 71

Variatio 19. a 1 Clav. · *Convexity forces metrizability* 74

Part III. Simplicity 77

Variatio 20. a 2 Clav. · *Simplicity* 78

Variatio 21. Canone alla Settima · *The Rado graph* 80

Variatio 22. a 1 Clav. Alla breve · *The Pauli group* 82

Variatio 23. a 2 Clav. · *The dictionary of simplicity* 85

Variatio 24. a 1 Clav. Canone all'Ottava · *Fraïssé limits and Clifford circuits* 87

Variatio 25. a 2 Clav. Adagio · *Forking and logical errors* 90

Variatio 26. a 2 Clav. · *The gluing of syndromes* 94

Variatio 27. a 2 Clav. Canone alla Nona · *The Rado graph inside the Pauli group* 96

Variatio 28. a 2 Clav. · *Mixing destroys simplicity* 98

Variatio 29. a 1 ô vero 2 Clav. · *Actual and idealized trees* 101

Variatio 30. a 1 Clav. Quodlibet · *The commutation relation* 104

Aria da capo e fine 106

Proofs 107

Questions 164

Historical remarks 165

Acknowledgments 167

Contributions and antecedents 168

Bibliography 173

Abstract

In this memoir we show that Shelah's dividing lines become, through bipartite observables, distinctions among quantum measurements. Ultrapowers, an instrument common to logic and operator theory, supply the account of idealization on which these distinctions rest. An observable is stable exactly when the two orders of idealization agree, and the idealized responses of a stable observable are then approximated uniformly by finite mixtures of actual preparations, with the probing system idealized first; the uniform form of stability controls adaptive discrimination and fixed-threshold prediction at every positive margin. NIP, the absence of the independence property, sorts physical orders, from the counting of photons to the causal order of a spacetime with one dimension of space, by the topology of their limiting responses, through Todorčević's trichotomy. Simplicity is at work in quantum error correction, where forking separates, among the nontrivial errors that no check detects, those that do no harm from those that do. Mixing, the operation that distinguishes quantum observables from arbitrary formulas, leaves its mark on each line, and on the last most sharply: for diagonal observables, mixing turns the independence property of the pure responses into the tree property of the second kind, so that a simple diagonal observable is NIP. For the Pauli group, actual mixtures of checks already witness this tree property, whereas for the random graph no quantifier-free formula in one probe variable has a tree property witnessed by actual preparations. We formulate the physical examples for the logician, and explain types and independence to the physicist and operator theorist. Both will find a single theorem of Grothendieck at the bottom of stability, and its broken form beneath NIP. A reader who belongs to neither community may also read the memoir, provided that ultrafilters hold no terrors, being less alarming than their name.

Introduction

In this memoir we show that the abstractions of logic, applied to quantum measurements, become instruments of physical description. If, in place of the formulas of a first-order theory, one considers bipartite observables, measurements performed on a pair of quantum systems, then the dividing lines that Shelah drew among theories become distinctions among measurements, and the theorems by which each line is characterized become assertions about what can be observed.

The operator theorist and the logician possess a common instrument: the ultrapower. In analysis it enlarges a Banach space or an operator algebra so that bounded sequences are represented by elements of a single structure; in model theory it is an elementary extension in which these elements have types over the original structure. The two descriptions direct attention to different features of the same construction. Types and independence, on which Shelah erected his classification of theories, can therefore be studied in the operator-theoretic setting without changing the underlying instrument, and the expectation of a bipartite observable supplies the connection with preparations and measurements. The idealized preparations of physics, a particle with a sharp momentum or an oscillator excited to infinite energy, appear as limits of actual preparations along an ultrafilter, and with two systems, Alice's and Bob's, the question arises in which order they are idealized.

Stability, the first of Shelah's lines, is the condition under which this order does not matter. An observable is stable exactly when the two orders of idealization agree, and for a stable observable every idealized state of Alice's system is then simulated by finite mixtures of actual preparations, uniformly over all probes, idealized ones included, with Bob's system idealized first; for the overlap measurements of quantum information, the number of terms needed does not depend on the dimension.

The second line, NIP, the absence of the independence property, is read in the topology of limiting responses. An observable is NIP exactly when its idealized responses are of the first Baire class, by the dichotomy of Rosenthal and of Bourgain, Fremlin and Talagrand. For an NIP observable, the compact space of these responses lies in one of three topological classes, because mixtures make it convex. Each class contains a physical order, from the counting of photons to the causal order of a spacetime with one dimension of space.

The third line, simplicity, is at work in quantum error correction. The Pauli group on countably many qubits, the two-level quantum systems, is generated by their bit flips and phase flips, and, modulo phases, it is the Fraïssé limit, in a sense recalled below, of the finite patterns of commutation among its elements, together with their linear relations. Among the errors other than the identity that no check detects, forking separates those that do no harm from those that do.

Mixing, the operation that distinguishes quantum observables from arbitrary formulas, runs through all three lines, and in the last it decides the matter: for diagonal observables, mixing turns the independence property into the tree property of the second kind, so that a simple diagonal observable is NIP. For the Pauli group, actual mixtures of checks already

witness this tree property, whereas for the random graph the witnesses of a quantifier-free formula in one probe variable must include idealized preparations.

Much of the work presented here grew out of conversations with Eduardo Dueñez, whose contributions to its development are acknowledged with gratitude, and here we gather the results into a single account, with proofs; the contributions of the memoir are listed, with their antecedents, in a table at the end (see Table 9).

The account takes the form of an aria with thirty variations. In the Aria we state Grothendieck's double limit exchange condition and explain its physical meaning, and at the end of the memoir the Aria returns, da capo, and is heard against what the three parts have established. The first variation, on observables as structures, is preparatory and closes with the conventions of model theory that the later variations recall; the other twenty-nine form three parts of parallel design, on stability, on NIP and on simplicity, in order that each part may be read against the others. Each part opens with the configuration that its line forbids, the theorem that characterizes the line, a dictionary in which the model theory of the line is rendered in physical terms, and the Fraïssé limits in which the line is exhibited: Hilbert space for stability, the rational order for NIP, and, for simplicity, the random graph and the Pauli group. Each of these structures is the Fraïssé limit of the class of its finitely generated substructures, as the rational order is the Fraïssé limit of the class of finite linear orders: it is countable or, like Hilbert space, separable and complete, and it is homogeneous, in the sense that every isomorphism between two finitely generated substructures extends to an automorphism. These properties, together with the class of its finitely generated substructures, determine it up to isomorphism. Near its end, each part contains a variation on mixing: states can be mixed, and the expectation of a bipartite observable at a mixture of states of either system is the weighted average of its expectations at the components. Many variations close with a paragraph for the physicist or operator theorist, one for the logician, and a sentence recording what has been established, and the proofs too long for the variations are collected after the Aria da capo, in a section of their own. In Table 1 we set the three parts side by side, and in the box below we state the main results; each notion used in the table and in the box is defined in a later variation.

	I. Stability	**II. NIP**	**III. Simplicity**
forbidden configuration	a ladder (order property)	an independent sequence	a tree whose branches are consistent and whose siblings are k-inconsistent, for some fixed integer $k \geq 2$
characterization	the two iterated limits agree (Grothendieck); an independence relation with stationarity over models (Harnik, Harrington)	the first Baire class (Rosenthal; Bourgain, Fremlin, Talagrand)	an independence relation with the independence theorem over models (Kim, Pillay)
nonforking extensions of a type over a model	unique	global ones invariant over the model, possibly several	possibly several, amalgamated by the independence theorem
Fraïssé limit	Hilbert space	the rational order	the random graph; the Pauli group
dictionary	orders of idealization	physical orders	quantum error correction
what mixing does	uniformly approximates idealized responses, with the probing system idealized first	makes hereditary separability imply metrizability	makes simplicity imply NIP for diagonal observables

TABLE 1. The three parts side by side.

Main results.

(I) *Stability.* An observable is stable exactly when the two orders of idealization agree. This condition has seven equivalent forms. They include uniform approximation of idealized responses by finite mixtures, with the probing system idealized first, and the agreement of the two orders in which the expectation types determined by a state of each system are realized; the common values, taken over all stable observables, form a symmetric independent product of the two states. At each stable observable, the expectation of this product is the unique separately weak* continuous extension of its values on pairs of actual states. A strictly stronger, uniform condition is equivalent to polynomial growth of adaptive expectation-response transcripts at every positive margin, with the threshold fixed within each protocol (see Theorems 5.1 and 7.1).

(II) *NIP.* An observable is NIP exactly when its idealized responses are of the first Baire class. Convexity, through the metrizability criterion of Todorčević and Gruenhage, leaves exactly three topological classes for the full response compacta of NIP observables: metrizable; first countable but not hereditarily separable; and not first countable. All three occur among physical orders, even under uniform NIP bounds at every positive margin. Each of the seven minimal families of the Argyros–Dodos–Kanellopoulos heptachotomy is realized by a tree of pure preparations, and in each nonmetrizable class every trace-norm dense sequence of preparations contains a tree subfamily that exhibits the appropriate minimal obstruction (see Theorems 11.1, 19.1, 15.1 and 16.1).

(III) *Simplicity.* The Pauli group on countably many qubits is the Fraïssé limit of the finite alternating spaces of labels, finite commutation patterns together with their linear relations, and the notions of simple theories translate into quantum error correction: among the undetectable errors other than the identity, the harmless ones are those whose types over the stabilizer labels fork over the empty set. For diagonal observables, simplicity implies NIP: if the pure responses have the independence property, then the mixtures carry the tree property of the second kind, and the observable is not simple. For the commutation observable of the Pauli group, actual mixtures already witness this tree property, whereas for the Rado observable no fixed quantifier-free formula in one probe variable has a tree property witnessed by actual preparations (see Variations 23–26 and Theorem 28.1 and Theorem 29.1).

In Figure 1 we place the observables of Parts I and II in the classes of stable and of NIP observables, together with the minimal families of trees that their preparations realize; for the Ising model, the ordered phase is the one with positive spontaneous magnetization, and the disordered phase the one with zero spontaneous magnetization.

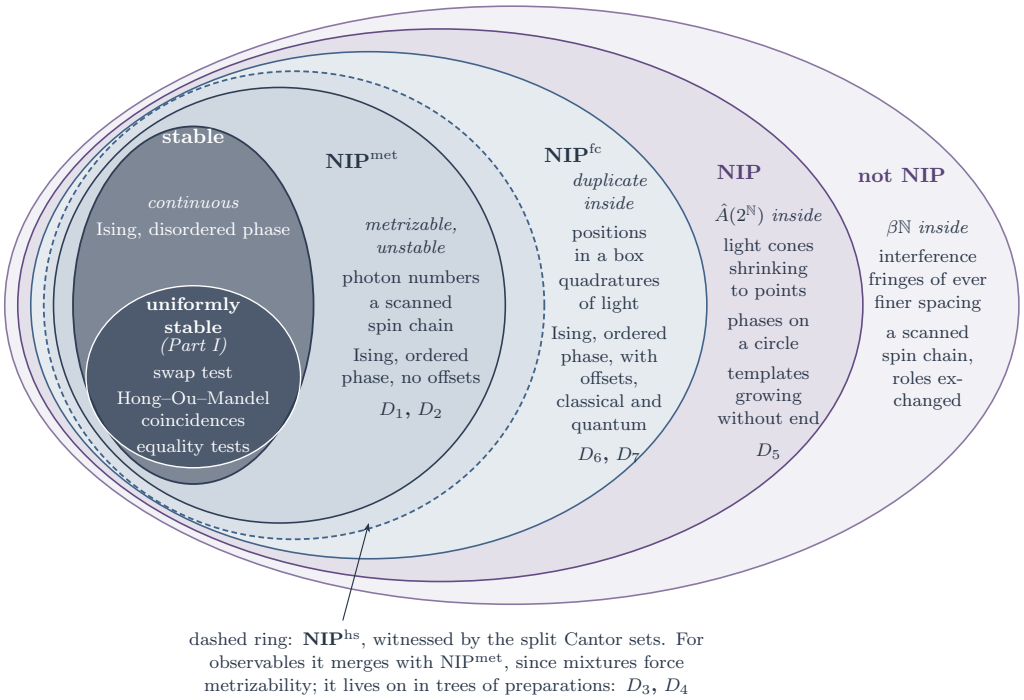


FIGURE 1. The classes of uniformly stable, stable and NIP observables, with the physical observables that inhabit them and the minimal families $D_1, \ldots, D_7$ realized by their trees of preparations.

Aria

Physics lives on idealized limits. A physicist speaks freely of systems that no laboratory has ever prepared: a particle with a sharp momentum, a particle localized at a point, an oscillator excited to infinite energy. None of them is produced by an actual preparation, but each is approached by sequences of actual preparations: wave functions concentrated ever more sharply in momentum or in position, states of ever higher energy. In the framework used here, an idealized preparation is modeled by a limiting state, obtained from a sequence of actual preparations once an ultrafilter is chosen; different sequences, or different ultrafilters, can give different limits that fit the same physical description.

To make this precise, let a quantum system be described by a Hilbert space H; its actual preparations are described by *density operators*, positive operators on H of trace 1 (see Variatio 1). The bounded operators on H form an algebra $B(H)$, and a *state* of $B(H)$ is a positive linear functional on $B(H)$ with value 1 at the identity; each density operator ρ gives the state $A \mapsto \mathrm{Tr}(\rho A)$, where Tr is the trace. Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$, and let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence of density operators on H. Recall that every bounded sequence $(a_n)_{n\in\mathbb{N}}$ of complex numbers has a unique ultralimit $\lim_{n,\mathcal{U}} a_n$, the complex number a such that $\{n : |a_n - a| < \varepsilon\} \in \mathcal{U}$ for every $\varepsilon > 0$; it is the ordinary limit whenever the latter exists. The *limit* of $(\rho_n)_{n\in\mathbb{N}}$ along $\mathcal{U}$ is the state $\lim_{n,\mathcal{U}} \rho_n$ of $B(H)$ that assigns to each $A \in B(H)$ the value $\lim_{n,\mathcal{U}} \mathrm{Tr}(\rho_n A)$; this state need not be given by any density operator. The energy eigenstates of an oscillator, its states of definite energy, form a sequence whose energies grow without bound (we describe the oscillator in the example below); along every nonprincipal ultrafilter on $\mathbb{N}$, the limit of this sequence is a state that is not given by any density operator and assigns probability zero to every finite set of energy levels. For a single system, idealization raises no question of order: by definition, the expected value of an observable, a bounded self-adjoint operator A on H, in the state $\lim_{n,\mathcal{U}} \rho_n$ is the ultralimit along $\mathcal{U}$ of its expected values $\mathrm{Tr}(\rho_n A)$ in the actual preparations.

Difficulties begin with two systems. Quantum information theory describes its procedures, called protocols, in terms of parties, each of whom controls a system in a separate laboratory; here there are two, by custom Alice and Bob. Alice's system is described by a Hilbert space H_1 and Bob's by a Hilbert space H_2; Alice and Bob prepare their systems independently and submit the pair to a joint measurement. The two systems together are described by the tensor product $H_1 \otimes H_2$, and a *bipartite observable* is a bounded self-adjoint operator C on $H_1 \otimes H_2$, which describes a measurement performed on the pair, with an outcome that depends jointly on both systems. When Alice prepares her system in the state given by a density operator ρ on H_1 and Bob, independently, prepares his in the state given by a density operator σ on H_2, the joint state is the one given by $\rho \otimes \sigma$, and the expected outcome of C is $\mathrm{Tr}\big((\rho\otimes\sigma)C\big)$. The two roles are symmetric in principle; we speak of Alice's preparations and Bob's probes because Alice's state is the object under study and Bob's states are the tests applied to it, much as a specimen is distinguished from the instruments that examine it.

A logician will recognize the division: a bipartite observable C defines a formula $\varphi(x, y)$ whose variables are partitioned, as Shelah partitions them, with Alice holding x and Bob holding y, and the number $\mathrm{Tr}\big((\rho \otimes \sigma)C\big)$ is the value of this formula at the pair (ρ, σ). In the same language, a sequence $(\rho_n)_{n\in\mathbb{N}}$ of Alice's density operators gives an element of the ultrapower along $\mathcal{U}$ of the structure that C defines (see Variatio 1), namely the class $[\rho_n]_{\mathcal{U}}$ of the sequence; the state $\lim_{n,\mathcal{U}} \rho_n$ of $B(H_1)$ determines the *φ-type* of this element, the function that records the value of φ at the pair formed by the element and each density operator of Bob. When both systems are idealized, Alice's along the sequence $(\rho_n)_{n\in\mathbb{N}}$ and Bob's along a sequence $(\sigma_m)_{m\in\mathbb{N}}$ of his probes, the expected outcome of C becomes one of two iterated limits,

$$\lim_n \lim_m \mathrm{Tr}\big((\rho_n \otimes \sigma_m)C\big) \quad \text{or} \quad \lim_m \lim_n \mathrm{Tr}\big((\rho_n \otimes \sigma_m)C\big),$$

the first with Bob's system idealized first and the second with Alice's, whenever these limits exist; nothing guarantees that the two agree.

A simple example is already decisive. An oscillator is described by the Hilbert space $L^2(\mathbb{R})$ of square-integrable functions of its position q and by its energy operator, or *Hamiltonian*, the unbounded self-adjoint operator $\frac{1}{2}(-d^2/dq^2 + q^2)$, written in units in which the reduced Planck constant, the mass and the angular frequency equal 1. The Hamiltonian has an orthonormal basis of eigenvectors $|0\rangle, |1\rangle, |2\rangle, \ldots$, the eigenvector $|k\rangle$ with eigenvalue $k + \frac{1}{2}$, and in the state given by $|k\rangle$ a measurement of energy gives $k + \frac{1}{2}$ with certainty. Only the basis and the order of the energies are used below: the measurements are built from the orthogonal projections onto the lines spanned by the eigenvectors, and these projections are bounded operators. Suppose that two identical oscillators, Alice's and Bob's, are prepared in energy eigenstates, Alice's in $|k\rangle_A$ and Bob's in $|l\rangle_B$, the subscripts telling the two oscillators apart, and that a measurement asks whether Bob's oscillator has the higher energy. The answer a_{kl} is 1 when $k < l$, and 0 when $k \geq l$. Then we have

$$\lim_l \lim_k a_{kl} = 0, \qquad \lim_k \lim_l a_{kl} = 1.$$

The matrix $(a_{kl})_{k,l\geq 0}$ has ones above the diagonal and zeros on and below it; we call this configuration the *staircase*. If Bob's oscillator is driven to infinite energy first, it outruns Alice's, and the answer is 1; if Alice's is driven there first, it outruns Bob's, and the answer is 0. Once both oscillators are idealized, these two procedures assign different probabilities to the event "Bob's oscillator has the higher energy": the value depends on which system is idealized first.

The agreement of the two orders has a finite approximation consequence. Fix a bipartite observable C, and suppose that, for all sequences of Alice's preparations and of Bob's probes, the two iterated limits of the expectation of C agree whenever both exist, that is, that C is stable (see Variatio 2). We show that every state of Alice's algebra is then simulated, to arbitrary accuracy in its responses and uniformly over all probes, idealized ones included, by finite mixtures of actual preparations, with Bob's system idealized first (see Theorem 5.1). For a state of each system, the common values also define an independent product of the two states on the space of stable observables; at each stable observable, the expectation of the product is the unique extension of the expectations at pairs of actual states that is weak* continuous in each marginal state separately. A strictly stronger condition, uniform stability (see Variatio 7), governs finite discrimination, in which Bob chooses each probe in view of the previous answers and records only whether the expectation lies above or below a threshold by at least a positive margin: at every positive margin, some polynomial bounds, for every positive integer N, the number of binary transcripts realized by N adaptive

expectation queries, and thresholded responses can be predicted online with a finite mistake bound, uniform in the threshold (see Theorem 7.1). The threshold is fixed within each protocol; estimating an expectation from repeated preparations is a separate statistical task. When uniform stability fails, there is a margin at which, for every positive integer N, all 2^N transcripts occur; qualitative stability alone need not prevent this exponential behavior.

The theme of this memoir is the theorem that says exactly when the two orders agree, and it was proved long before anyone asked the question in these terms. Grothendieck proved it in 1952, in his work on compactness in function spaces [**Gro52**]: let Y be a compact space, and let F be a bounded set of continuous functions on Y. Then every function in the pointwise closure of F is continuous if and only if, for all sequences $(g_n)_{n\in\mathbb{N}}$ in F and $(y_m)_{m\in\mathbb{N}}$ in Y, we have

$$\lim_n \lim_m g_n(y_m) = \lim_m \lim_n g_n(y_m)$$

whenever both iterated limits exist. Since F is bounded, the first condition is equivalent to the relative weak compactness of F in the Banach space $C(Y)$ of continuous functions on Y with the supremum norm; a topological property of the whole set is thus decided by the arrays $\big(g_n(y_m)\big)_{n,m\in\mathbb{N}}$ of its values along sequences.

For Y, take the set of all states of $B(H_2)$, actual or idealized, which is compact in the weak* topology, the topology of convergence at every operator; for F, take the responses of Alice's actual preparations. Once Alice has prepared a density operator ρ on H_1, the observable C acts on Bob's system through the bounded operator $T_C(\rho)$ on H_2 with $\operatorname{Tr}(\sigma\, T_C(\rho)) = \operatorname{Tr}\big((\rho\otimes\sigma)C\big)$ for every density operator σ on H_2, and the *response* of ρ is the function $\omega \mapsto \omega(T_C(\rho))$ on Y. We call the functions in the pointwise closure of these responses the *idealized responses* of Alice, and the theorem says that a bipartite observable is insensitive to the order of idealization exactly when the idealized responses of Alice are continuous on Y.

A logician recognizes the staircase at once: it is an instance of the order property, the configuration whose absence defines the first of the dividing lines that Shelah drew across the first-order theories. Read for a formula $\varphi(x, y)$, Grothendieck's theorem says that φ is stable in a theory exactly when every φ-type over every model of the theory is *definable*, that is, given at each parameter b of the model by the value at b of a single formula, or of a uniform limit of formulas, with parameters from the model [**BY14**] (theories, models and types are defined in Variatio 1). The second line, NIP, separates the failures of the double limit condition in which idealized responses remain limits of sequences of responses of actual preparations from those in which they do not (see Theorem 11.1), and the third, simplicity, is characterized not by limits but by the amalgamation of independent types, through the Kim–Pillay theorem.

In Part I we study the observables that satisfy Grothendieck's double limit condition, in Part II the observables that are NIP, and in Part III the amalgamation of independent types, which in simple theories takes the place of the uniqueness of nonforking extensions.

VARIATIO 1

a 1 Clav.

Observables as structures

The two iterated limits of the staircase of the Aria, the matrix of answers to the question whether Bob's oscillator has the higher energy, idealize two systems at once, and so far they have been computed only along sequences of energy eigenstates. To compare the two orders of idealization for an arbitrary observable, the idealized systems must become objects in their own right, and the value that each order assigns to the observable must depend only on these objects, not on the sequences that approach them. In this variation we distinguish actual and idealized states of the algebra of bounded operators on a Hilbert space and define the two products that correspond to the two orders of idealization, in which Alice's system or Bob's is idealized first. We then attach to each bipartite observable a real-valued structure, show that, once the observable is fixed, an idealized state determines the restriction to a single formula of the type of an element of an ultrapower of the structure, and obtain the two products from the two orders in which an iterated ultrapower realizes types.

States and observables

A quantum system is described by a complex separable Hilbert space H. Inner products are conjugate-linear in the first variable and linear in the second; we write $\langle\chi|\xi\rangle = \langle\chi, \xi\rangle$. For vectors $\psi, \chi \in H$, we write $|\psi\rangle\langle\chi|$ for the operator $\xi \mapsto \langle\chi, \xi\rangle\psi$. A unit vector $\psi \in H$ describes a *pure* preparation. For a positive operator T on H, the sum $\sum_n \langle\zeta_n, T\zeta_n\rangle$ over an orthonormal basis $(\zeta_n)_{n<\dim H}$ of H does not depend on the basis; when it is finite, the operator T is of *trace class*, and the sum is its *trace* $\operatorname{Tr} T$. The trace-class operators are the linear combinations of positive operators of trace class, and the trace extends linearly to them. A general preparation is described by a *density operator*, a positive trace-class operator ρ of trace 1; the pure preparations correspond to the rank-one orthogonal projections $|\psi\rangle\langle\psi|$, and a unit vector ψ also stands for the density operator $|\psi\rangle\langle\psi|$. The density operators form a convex set $\mathfrak{D}(H)$, and a *mixture* $\sum_{i=1}^k p_i\rho_i$ of density operators $\rho_1, \dots, \rho_k$, with $p_i \geq 0$ and $\sum_{i=1}^k p_i = 1$, describes a preparation that produces ρ_i with probability p_i. A projective measurement with finitely many outcomes, labeled $1, \dots, r$, is described by a family of orthogonal projections $E_1, \dots, E_r$ with sum I, the identity operator, and in the state given by ρ the outcome j occurs with probability $\operatorname{Tr}(\rho E_j)$, by the Born rule. If each outcome j is recorded as a real number λ_j, the measurement is represented by the operator $\sum_j \lambda_j E_j$, whose expectation $\operatorname{Tr}\bigl(\rho \sum_j \lambda_j E_j\bigr)$ in the state given by ρ is the expected value of the recorded number; more generally, we call every bounded self-adjoint operator A on H an *observable*: by the spectral theorem, it is a norm limit of real linear combinations of commuting orthogonal projections, and the number $\operatorname{Tr}(\rho A)$ is the expected value of its outcome in the state given by ρ. The expectation is affine in ρ: for all density operators ρ and ρ' and every t with $0 \leq t \leq 1$, we have $\operatorname{Tr}\bigl((t\rho + (1-t)\rho')A\bigr) = t\operatorname{Tr}(\rho A) + (1-t)\operatorname{Tr}(\rho' A)$, that is, the expected outcome of

a mixture is the weighted average of the expected outcomes of its components, as befits a preparation that chooses its components at random.

The systems of Alice and of Bob, the two experimenters of the Aria, are described by Hilbert spaces H_1 and H_2, and together by the Hilbert tensor product $H_1 \otimes H_2$, the completion of the algebraic tensor product under the inner product with $\langle \psi \otimes \chi, \psi' \otimes \chi' \rangle = \langle \psi, \psi' \rangle \langle \chi, \chi' \rangle$. When they are prepared independently, in the states given by density operators ρ on H_1 and σ on H_2, the joint state is the one given by $\rho \otimes \sigma$, and a bounded self-adjoint operator C on $H_1 \otimes H_2$, a *bipartite observable*, has expectation

$$\varphi_C(\rho, \sigma) = \operatorname{Tr}\big((\rho \otimes \sigma)C\big).$$

For Alice's preparation ρ, the *effective observable* $T_C(\rho)$ is the unique bounded operator on H_2 with $\operatorname{Tr}(\sigma\, T_C(\rho)) = \varphi_C(\rho, \sigma)$ for every density operator σ on H_2; it is obtained by the *partial trace* over Alice's system, $T_C(\rho) = \operatorname{Tr}_1\big((\rho \otimes I)C\big)$, defined by $\langle \eta, T_C(\rho)\xi \rangle = \sum_n \big\langle \zeta_n \otimes \eta, (\rho \otimes I)C(\zeta_n \otimes \xi) \big\rangle$ for all $\xi, \eta \in H_2$ and any orthonormal basis $(\zeta_n)_{n<\dim H_1}$ of H_1. For $\rho \in \mathfrak{D}(H_1)$, define the *response on actual probes* $r_\rho \colon \mathfrak{D}(H_2) \to [-\|C\|, \|C\|]$ by $r_\rho(\sigma) = \varphi_C(\rho, \sigma)$. The response of ρ in the Aria is the extension $\omega_2 \mapsto \omega_2(T_C(\rho))$ of r_ρ to all states ω_2 of $B(H_2)$, whose value at the state given by a density operator σ is $r_\rho(\sigma)$. In Variatio 11, we replace $B(H_2)$ by the C*-algebra generated by the effective observables and the identity, the probe algebra, and write $\mathfrak{r}_\rho$ for the completed response of ρ, a function on the state space of this algebra. For the energy comparison of the Aria, the measurement that asks whether Bob's oscillator has the higher energy, let $|0\rangle_A, |1\rangle_A, \dots$ and $|0\rangle_B, |1\rangle_B, \dots$ be the energy eigenvectors of Alice's oscillator and of Bob's; the observable is the orthogonal projection P below, and its effective observable at $|k\rangle_A$, for $k = 0, 1, 2, \dots$, is the orthogonal projection $T_P(|k\rangle_A)$:

$$P = \sum_{i<j} |i\rangle_A\langle i|_A \otimes |j\rangle_B\langle j|_B, \qquad T_P(|k\rangle_A) = \sum_{l>k} |l\rangle_B\langle l|_B.$$

Idealized states

A *state* of a system described by a Hilbert space H is a positive linear functional ω on $B(H)$ with $\omega(I) = 1$, as in the Aria. Each density operator ρ defines the *actual* state $A \mapsto \operatorname{Tr}(\rho A)$, which physicists call *normal*. We call the remaining states *idealized*. The states of $B(H)$ form a compact convex set in the weak* topology, the topology of pointwise convergence on operators, and the actual states are dense in this set. In this topology the energy eigenvectors $|0\rangle_A, |1\rangle_A, \dots$ of the oscillator of the Aria form a sequence with no limit, and the cluster points of the sequence are idealized states that are not actual. For a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the state $\omega = \lim_{k,\mathcal{U}} |k\rangle_A$, given by $\omega(A) = \lim_{k,\mathcal{U}} {}_A\langle k|A|k\rangle_A$, vanishes on every compact operator, that is, on every norm limit of operators of finite rank, such as the orthogonal projection onto the span of finitely many of the vectors $|k\rangle_A$; it therefore assigns probability zero to every finite set of energy levels. Such a state represents escape beyond every finite energy cutoff, and different ultrafilters give different such states.

Two orders of idealization

To compare the two orders of idealization, let ω_1 be a state of Alice's algebra $B(H_1)$, and ω_2 a state of Bob's algebra $B(H_2)$, and let $(\rho_i)_{i \in \mathcal{I}}$ and $(\sigma_j)_{j \in \mathcal{J}}$ be nets of density operators, families indexed by directed sets $\mathcal{I}$ and $\mathcal{J}$, that converge in the weak* topology to ω_1 and to ω_2, respectively; nets are used because a sequence of density operators that converges in this topology has a density operator as its limit, and therefore no idealized state that is not actual

is the limit of a sequence of density operators. The expectation of C can be idealized in two orders, which define two *products* of ω_1 and ω_2 at C:

$$(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = \lim_j \lim_i \varphi_C(\rho_i, \sigma_j), \qquad (\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \lim_i \lim_j \varphi_C(\rho_i, \sigma_j),$$

the first with Alice's system idealized first and the second with Bob's. Both limits exist and do not depend on the nets: for the first, the inner limit exists because, for each fixed density operator σ on H_2, the value $\varphi_C(\rho, \sigma)$ is $\operatorname{Tr}(\rho S)$ for the bounded operator $S = \operatorname{Tr}_2\big((I \otimes \sigma)C\big)$ on H_1, the partial trace over Bob's system, defined as Tr_1 above with the roles of the two factors exchanged, and the net $(\rho_i)_{i\in\mathcal{I}}$ converges to ω_1; the map $\sigma \mapsto \lim_i \varphi_C(\rho_i, \sigma)$ extends to a bounded linear functional on the trace-class operators on H_2, and it therefore has the form $\sigma \mapsto \operatorname{Tr}(\sigma Z)$ for a unique bounded operator $Z = Z_{\omega_1}$ on H_2, whence the first product is $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = \omega_2(Z_{\omega_1})$. The argument for the second product is symmetric. The two products are the two Arens extensions [**Are51**] of the bounded bilinear form $(X, Y) \mapsto \operatorname{Tr}\big((X \otimes Y)C\big)$ on pairs of trace-class operators X on H_1 and Y on H_2, which extends φ_C, that is, the two extensions of this form to the biduals of the two spaces of trace-class operators, which contain the states of $B(H_1)$ and of $B(H_2)$, obtained by taking weak* limits first in one variable and then in the other.

If one of the two states is actual, the two products agree: if ω_1 is given by a density operator ρ, then the operator Z_{ω_1} is $T_C(\rho)$, and both products equal $\omega_2(T_C(\rho))$; the case of ω_2 is symmetric. Disagreement needs two idealized systems.

For the energy comparison P, fix a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, and let $\omega_1 = \lim_{k,\mathcal{U}} |k\rangle_A$ and $\omega_2 = \lim_{l,\mathcal{U}} |l\rangle_B$. For each fixed l, the value $\varphi_P(|k\rangle_A, |l\rangle_B)$ is 0 for all $k \geq l$; for each fixed k, it is 1 for all $l > k$. Since the products do not depend on the nets, we may compute them along subnets of $(|k\rangle_A)_{k\in\mathbb{N}}$ and of $(|l\rangle_B)_{l\in\mathbb{N}}$ that converge to ω_1 and to ω_2, and along such subnets the indices tend to infinity; hence we have

$$(\omega_1 \otimes_{\mathsf{A}} \omega_2)(P) = 0, \qquad (\omega_1 \otimes_{\mathsf{B}} \omega_2)(P) = 1.$$

The two orders of idealization, with Alice's system and with Bob's system idealized first, assign to the event "Bob's oscillator has the higher energy" the probabilities 0 and 1, respectively.

The structure of an observable

The expectation of a bipartite observable is a bounded real-valued function of a preparation of Alice and a probe of Bob, and we treat the expectation as a formula of real-valued logic. We work in Keisler's model theory for real-valued structures [**Kei23**], in which a *general structure* assigns to each sort of a vocabulary a nonempty set and to each predicate symbol a function with values in the bounded interval of the symbol, with no distinguished metric and no prescribed moduli of uniform continuity; in the continuous logic of Ben Yaacov, Berenstein, Henson and Usvyatsov [**BYBHU08**], by contrast, every sort is a complete metric space and every predicate is uniformly continuous. A bipartite observable C determines a general structure $\mathcal{M}_C$ with two sorts, the sets $\mathfrak{D}(H_1)$ and $\mathfrak{D}(H_2)$ of density operators of Alice and of Bob, and three predicates: the expectation φ_C, with values in $[-\|C\|, \|C\|]$, and, on each sort, the trace-norm distance $(\rho, \rho') \mapsto \|\rho - \rho'\|_1$, with values in $[0, 2]$, where $\|T\|_1 = \operatorname{Tr}|T|$, with $|T| = (T^*T)^{1/2}$, is the trace norm of a trace-class operator T. We write φ for the predicate symbol of the expectation and φ_C for its interpretation in $\mathcal{M}_C$. Formulas are built from the atomic formulas with continuous connectives $u\colon \mathbb{R}^n \to \mathbb{R}$ and the quantifiers $\sup_x$ and $\inf_x$. For instance, the value of the formula $\sup_y |\varphi(x, y) - \varphi(x', y)|$ at a pair (ρ, ρ') of Alice's

density operators is the supremum of the differences that a probe of Bob can detect between the expectations of C under ρ and under ρ'. When the vocabulary is countable, Keisler's expansion theorem shows that every theory, a set of conditions on the values of sentences, has an expansion by a distance predicate, a pseudometric that is a uniform limit of formulas, and that this predicate is unique up to uniform equivalence; the structure $\mathcal{M}_C$ needs no such expansion, since its trace-norm distances are already predicates.

An ultrapower of $\mathcal{M}_C$ turns sequences of density operators, from which the idealized states of the Aria were obtained as limits along an ultrafilter, into elements of a single structure, at which formulas can be evaluated. Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$, and let $\mathcal{M}_C^{\mathcal{U}}$ be the ultrapower of $\mathcal{M}_C$ along $\mathcal{U}$, in which two sequences are identified when every atomic formula has the same ultralimit at both; we write $[\rho_n]_{\mathcal{U}}$ for the class of a sequence $(\rho_n)_{n\in\mathbb{N}}$. Since the trace-norm distance is a predicate and the expectation is Lipschitz in each argument with respect to it, with constant $\|C\|$, two sequences $(\rho_n)_{n\in\mathbb{N}}$ and $(\rho'_n)_{n\in\mathbb{N}}$ are identified exactly when $\lim_{n,\mathcal{U}} \|\rho_n - \rho'_n\|_1 = 0$, and the sorts are therefore the metric ultrapowers of $\mathfrak{D}(H_1)$ and $\mathfrak{D}(H_2)$. Measurements detect half the trace-norm distance: the largest difference between the probabilities that a single measurement assigns to an outcome under ρ and under ρ' is $\frac{1}{2}\|\rho - \rho'\|_1$, and two sequences of preparations are therefore identified in $\mathcal{M}_C^{\mathcal{U}}$ exactly when no sequence of measurements tells them apart along $\mathcal{U}$. The operator norm would identify too much: for each $n \geq 1$, let G_n and G'_n be orthogonal projections of rank n with orthogonal ranges; the operator norm of $G_n/n - G'_n/n$ is $1/n$, whereas its trace norm is 2, and the measurement given by G_n and $I - G_n$ tells the preparations G_n/n and G'_n/n apart with certainty. The operator norm would thus identify the sequences $(G_n/n)_{n\geq 1}$ and $(G'_n/n)_{n\geq 1}$, which the trace-norm distance keeps apart. Each sort is thus a closed convex subset of the Banach ultrapower, in the sense of Heinrich [**Hei80**], of the space of trace-class operators.

On the side of observables, the corresponding construction is the Groh–Raynaud ultrapower of $B(H)$, whose predual is the Banach ultrapower of the space of trace-class operators [**Gro84**, **Ray02**].

A formula $\varphi(x, y)$ is *stable in a theory* if no model of the theory has an infinite *ladder* for it, that is, elements $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ of the sorts of x and of y, together with reals t and $\delta > 0$, such that the value of φ at (a_i, b_j) is at most $t - \delta/2$ for $i \leq j$ and at least $t + \delta/2$ for $i > j$, or the same with the two inequalities exchanged (see Variatio 2). For an element a of an elementary extension $\mathcal{N}$ of a structure $\mathcal{M}$, the *φ-type* of a over $\mathcal{M}$ is the restriction of the type of a over $\mathcal{M}$ to φ, the function $b \mapsto \varphi^{\mathcal{N}}(a, b)$ on $\mathcal{M}$, where $\theta^{\mathcal{N}}$ denotes the interpretation of a formula θ in $\mathcal{N}$; in this notation, $\varphi^{\mathcal{M}_C}$ is φ_C.

Ultrapowers of $\mathcal{M}_C$, and ultrapowers of these, are the most concrete elementary extensions, and they suffice for the arguments of Parts I and II. Since $\mathcal{U}$ is nonprincipal, and hence countably incomplete (some countable family of its members has empty intersection), and the vocabulary is countable, the ultrapower $\mathcal{M}_C^{\mathcal{U}}$ is $\aleph_1$-saturated; since the sorts of $\mathcal{M}_C$ are, moreover, separable in the trace-norm distance, every elementary extension of $\mathcal{M}_C$ with separable sorts embeds elementarily in $\mathcal{M}_C^{\mathcal{U}}$ by an embedding that extends the diagonal map [**BYBHU08**]. Every state of $B(H)$ is the limit, along some ultrafilter on $\mathbb{N}$, of a sequence of density operators, because the density operators form a subset of the states that is separable in the trace norm and dense in the weak* topology. The ultrafilter must depend on the state: for a fixed $\mathcal{U}$ there are only $2^{\aleph_0}$ sequences of density operators, whereas $B(H)$ has $2^{2^{\aleph_0}}$ states when H is infinite-dimensional. Sequences thus reach every idealized state once the ultrafilter is allowed to vary. Part III needs a larger elementary extension, the monster model recalled

at the end of this variation, an elementary extension that realizes every type over every set of its elements that is small compared with the model.

The φ-types of elements of the ultrapower come from idealized states. Let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence of Alice's density operators, and let $\omega = \lim_{n,\mathcal{U}} \rho_n$ be its limit along $\mathcal{U}$, the state of $B(H_1)$ defined in the Aria. The φ-type over $\mathcal{M}_C$ of the element $a = [\rho_n]_{\mathcal{U}}$ of Alice's sort in $\mathcal{M}_C^{\mathcal{U}}$ is the function

$$\sigma \longmapsto \lim_{n,\mathcal{U}} \operatorname{Tr}\big((\rho_n \otimes \sigma)C\big)$$

on Bob's density operators; this function is $\sigma \mapsto \omega\big(\operatorname{Tr}_2((I\otimes\sigma)C)\big)$, and we call it the *response of ω on actual probes*. Thus, for a fixed observable C, the state ω determines the φ-type over $\mathcal{M}_C$ of every element $[\rho_n]_{\mathcal{U}}$ with $\lim_{n,\mathcal{U}} \rho_n = \omega$; different states may determine the same φ-type, and the full type of $[\rho_n]_{\mathcal{U}}$ is not determined by ω, since the full type also records the trace-norm distances from $[\rho_n]_{\mathcal{U}}$ to the elements of $\mathcal{M}_C$. Because formulas are uniformly continuous in the trace-norm distances, a type over $\mathcal{M}_C$ is determined by its values at the tuples from a countable dense subset of $\mathcal{M}_C$; since $\mathcal{M}_C^{\mathcal{U}}$ is $\aleph_1$-saturated, every type over $\mathcal{M}_C$ is therefore realized in $\mathcal{M}_C^{\mathcal{U}}$.

Independence enters with two systems. Let $\mathcal{V}$ be a nonprincipal ultrafilter on $\mathbb{N}$, possibly equal to $\mathcal{U}$, and let $(\rho_n)_{n\in\mathbb{N}}$ and $(\sigma_m)_{m\in\mathbb{N}}$ be sequences of Alice's and of Bob's density operators; their classes $a = [\rho_n]_{\mathcal{U}}$ and $b = [\sigma_m]_{\mathcal{V}}$ are elements of Alice's sort in $\mathcal{M}_C^{\mathcal{U}}$ and of Bob's sort in $\mathcal{M}_C^{\mathcal{V}}$, respectively. Before φ can be evaluated at the pair (a, b), the two elements must be placed in a common elementary extension of $\mathcal{M}_C$. When $\mathcal{U} = \mathcal{V}$, placing both in $\mathcal{M}_C^{\mathcal{U}}$ yields the joint limit $\lim_{n,\mathcal{U}} \varphi_C(\rho_n, \sigma_n)$, which need not equal either iterated limit and need not make a and b independent (see Variatio 6); for the energy comparison P it is 0, the value with Alice's system idealized first, when $\rho_n = |n\rangle_A$ and $\sigma_n = |n\rangle_B$, and 1, the value with Bob's system idealized first, when $\rho_n = |n\rangle_A$ and $\sigma_n = |n+1\rangle_B$. An iterated ultrapower offers two other ways of placing the two elements together. If a is taken in $\mathcal{M}_C^{\mathcal{U}}$ and the type of b over $\mathcal{M}_C$ is then realized in $(\mathcal{M}_C^{\mathcal{U}})^{\mathcal{V}}$ by the class of $(\sigma_m)_{m\in\mathbb{N}}$, the formula φ takes at the pair the value $\lim_{m,\mathcal{V}} \lim_{n,\mathcal{U}} \varphi_C(\rho_n, \sigma_m)$; in the other order it takes the value $\lim_{n,\mathcal{U}} \lim_{m,\mathcal{V}} \varphi_C(\rho_n, \sigma_m)$. These two values are the two products defined above: if ω_1 and ω_2 are the limits of $(\rho_n)_{n\in\mathbb{N}}$ along $\mathcal{U}$ and of $(\sigma_m)_{m\in\mathbb{N}}$ along $\mathcal{V}$, then the first value, in which Alice's element is placed first and her system is thus idealized first, is $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C)$, and the second is $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C)$. In the first order, the type over $\mathcal{M}_C^{\mathcal{U}}$ of the class of $(\sigma_m)_{m\in\mathbb{N}}$ is the limit along $\mathcal{V}$ of the types over $\mathcal{M}_C^{\mathcal{U}}$ of the elements σ_m of $\mathcal{M}_C$, and is therefore a coheir of the type of b over $\mathcal{M}_C$.

Conventions of model theory

We collect here the conventions of model theory that later variations use; each is recalled briefly where it is first needed. The *theory* of a structure records the value of every sentence in the structure, and a theory is *complete* if it is the theory of a structure. A *monster model* of a complete theory T is a model $\mathbb{M}$ of T that is κ-saturated and strongly κ-homogeneous for a very large cardinal κ: every type over a set of fewer than κ elements, a *small* set, is realized in $\mathbb{M}$, and two small tuples with the same type are conjugate by an automorphism of $\mathbb{M}$. The ultrapowers along nonprincipal ultrafilters on $\mathbb{N}$ used above have the first property for countable sets of parameters, when the language is countable. We write $\operatorname{tp}(a/B)$ for the type of a over B, and call a type over the whole monster model *global*. A sequence $(b_i)_{i\in\mathbb{N}}$ is *indiscernible* over a set A if, for every positive integer n, any two increasing n-tuples of its terms have the same type over A. A formula $\theta(x, b)$ *divides* over A if, for some integer

k, there is a sequence $b = b_0, b_1, b_2, \dots$ indiscernible over A that is *k-inconsistent* for θ: no k of the formulas $\theta(x, b_i)$ have a common solution; in real-valued logic, where $\theta(x, b)$ is read as the condition $\theta(x, b) = 0$, k-inconsistency is required with a margin: for some $\varepsilon > 0$, no x satisfies $|\theta(x, b_i)| \leq \varepsilon$ for k distinct indices i. A type *forks* over A if the type implies a finite disjunction of formulas that divide over A, and we write $a \mathop{\smile\!\!\!\!\!\!|}_A b$ when $\mathrm{tp}(a/Ab)$, the type of a over $A \cup \{b\}$, does not fork over A. For the classical, two-valued theories of Part III, a theory *eliminates quantifiers* if every formula is equivalent in the theory to a formula without quantifiers. In such a theory, an element a of a monster model is *algebraic* over a small set A if $\mathrm{tp}(a/A)$ has only finitely many realizations, and the *algebraic closure* of A is the set of such elements.

For the physicist or operator theorist. The sorts of the ultrapower $\mathcal{M}_C^{\mathcal{U}}$ are the metric ultrapowers of $\mathfrak{D}(H_1)$ and $\mathfrak{D}(H_2)$, closed convex subsets of the Banach ultrapowers of the spaces of trace-class operators. In logic, an element $[\rho_n]_{\mathcal{U}}$ of Alice's sort has a *type*, the record of the values of all formulas at the element; the restriction of this type to the expectation is the response, on actual probes, of the state $\lim_{n,\mathcal{U}} \rho_n$. The two products $\otimes_{\mathsf{A}}$ and $\otimes_{\mathsf{B}}$ are the values of the expectation at a pair of such elements placed together in an iterated ultrapower, in the two possible orders, and these placements are the independent amalgams of Variatio 6.

For the logician. The number $\varphi_C(\rho, \sigma)$ is the expected outcome of the observable C when Alice's system is prepared in ρ and Bob's, independently, in σ. An idealized state, such as the state $\lim_{k,\mathcal{U}} |k\rangle_A$ of the oscillator excited to infinite energy, is given by no density operator, and no laboratory prepares it. The two products $\otimes_{\mathsf{A}}$ and $\otimes_{\mathsf{B}}$, one for each order of idealization, are given by the two Arens extensions of the bilinear form $(X, Y) \mapsto \mathrm{Tr}\big((X \otimes Y)C\big)$ on trace-class operators, which extends φ_C; for the energy comparison they assign to the event "Bob's oscillator has the higher energy" the probabilities 0 and 1, respectively. What distinguishes observables from arbitrary formulas is that states can be mixed, and that the expectation of C at a mixture, whether of Alice's preparations or of Bob's, is the weighted average of its expectations at the components.

In Theorem 3.1 below, we show that Grothendieck's double limit condition, read for the expectation of C, is the statement that the two products agree for every pair of states. The staircase of the oscillators fails it with a gap of 1, the largest that an orthogonal projection allows.

Part I. Stability

VARIATIO 2

a 1 Clav.

The order property

The energy of a quantum system is described by a self-adjoint operator, the *Hamiltonian*; when the Hamiltonian has an orthonormal basis of eigenvectors, as for the two oscillators of the Aria, each eigenvalue is a possible energy, and each unit eigenvector describes a state of definite energy. Two such systems, one in Alice's laboratory and the other in Bob's, can be compared by two measurements: the first asks whether the two systems are in levels of the same energy, and the second asks whether Bob's system has the higher energy. Both measurements are classical, in the sense that they commute with each other and with every measurement of energy on either system; nevertheless, they behave in opposite ways once the systems are excited without bound. For the two oscillators, the second measurement is the energy comparison of the Aria. Its staircase is the prototype of the configuration forbidden by stability. Stability, the first of Shelah's dividing lines, is the subject of Part I. In this variation we define ladders and stability, show that a bipartite observable is unstable exactly when it has an infinite ladder (see Theorem 2.1), and explain the opposite behavior of the two measurements (see Proposition 2.2).

To write the two measurements as operators, let A and B be the Hamiltonians of Alice's system and of Bob's, and, for each real λ, let Π^A_λ and Π^B_λ be the orthogonal projections onto the eigenspaces of A and of B for λ, which are zero unless λ is an eigenvalue. The two measurements are then the orthogonal projections

$$[A = B] = \sum_\lambda \Pi^A_\lambda \otimes \Pi^B_\lambda, \qquad [A < B] = \sum_{\lambda<\lambda'} \Pi^A_\lambda \otimes \Pi^B_{\lambda'},$$

in which only countably many terms are nonzero; we call them the *equality test* and the *order test*.

Ladders

The definition of a ladder copies the staircase of the Aria up to a margin: the values of the expectation of an observable need not be 0 and 1, but they lie on one side of a threshold or on the other according to the position of the entry relative to the diagonal; a fixed margin keeps the two sides apart. Let C be a bipartite observable with expectation $\varphi_C(\rho, \sigma) = \mathrm{Tr}((\rho\otimes\sigma)C)$, and let $\delta > 0$ be given. A *ladder at margin* δ consists of a real threshold t and of density operators ρ_i of Alice and σ_j of Bob, for $1 \le i, j \le N$, where N is a positive integer, or for all positive integers i and j, such that

$$\varphi_C(\rho_i, \sigma_j) \le t - \tfrac{\delta}{2} \text{ for } i \le j, \qquad \varphi_C(\rho_i, \sigma_j) \ge t + \tfrac{\delta}{2} \text{ for } i > j,$$

or the same with the two inequalities exchanged; when the indices range over $1, \ldots, N$, the ladder is finite, of *length* N, and when they range over all positive integers, it is infinite. Ladders for a formula $\varphi(x, y)$ in a structure are defined in the same way, with elements of the two sorts in place of density operators; the ladders of C are the ladders of the formula φ

in the structure $\mathcal{M}_C$ of Variatio 1, whose two sorts are the sets of density operators of Alice and of Bob. A formula has Shelah's *order property* [She71] in a theory if, at some margin, it has ladders of every finite length in some model of the theory; by Łoś's theorem, ladders of every finite length at a fixed margin yield an infinite ladder at that margin in an ultrapower. A formula is therefore stable in a theory, as defined in Variatio 1, exactly when the formula does not have the order property in that theory. The observable C is *stable* if it satisfies Grothendieck's double limit condition: for all sequences $(\rho_k)_{k\in\mathbb{N}}$ of Alice's density operators and $(\sigma_l)_{l\in\mathbb{N}}$ of Bob's, we have

$$\lim_l \lim_k \varphi_C(\rho_k, \sigma_l) = \lim_k \lim_l \varphi_C(\rho_k, \sigma_l)$$

whenever both iterated limits exist.

In Figure 2 we draw as a ladder the staircase of the Aria, the matrix of values of the energy comparison of two oscillators. Of the two implications in Theorem 2.1 below, the one from instability to an infinite ladder carries the content: a disagreement between two iterated limits is a statement about limits alone, and the proof must extract from it a combinatorial configuration.

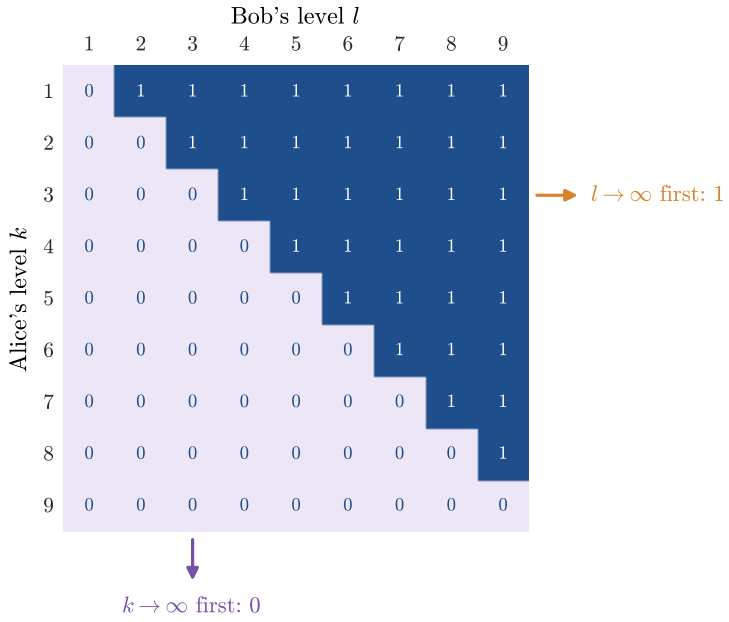


FIGURE 2. The values $\varphi_P(|k\rangle_A, |l\rangle_B)$ of the energy comparison P of two oscillators, in row k and column l.

Theorem 2.1. *An observable is unstable if and only if it has an infinite ladder at some margin.*

PROOF. Given an infinite ladder $(\rho_i)_{i\geq 1}$, $(\sigma_j)_{j\geq 1}$ at margin δ with threshold t, consider the bounded matrix $\big(\varphi_C(\rho_i, \sigma_j)\big)_{i,j\geq 1}$. A diagonal argument yields subsequences of its rows and of its columns along which both iterated limits exist; these limits lie on opposite sides of t, at distance at least δ from each other, and the observable C is therefore unstable.

Conversely, suppose that C is unstable: there are sequences $(\rho_k)_{k\in\mathbb{N}}$ of Alice's density operators and $(\sigma_l)_{l\in\mathbb{N}}$ of Bob's for which the matrix of values $a_{kl} = \varphi_C(\rho_k, \sigma_l)$ has iterated limits $L_1 = \lim_k \lim_l a_{kl}$ and $L_2 = \lim_l \lim_k a_{kl}$ with $L_1 \neq L_2$. The idea is to choose rows and

columns alternately, so that each entry of the resulting grid is close to the limit of its row or to that of its column, according to its position relative to the diagonal. Let $\alpha_k = \lim_l a_{kl}$ and $\beta_l = \lim_k a_{kl}$, whence $\alpha_k \to L_1$ as $k \to \infty$ and $\beta_l \to L_2$ as $l \to \infty$, and fix $\varepsilon > 0$ with $4\varepsilon < |L_1 - L_2|$. Having chosen $k_1 < \cdots < k_{m-1}$ and $l_1 < \cdots < l_{m-1}$, choose $k_m > k_{m-1}$ with $|\alpha_{k_m} - L_1| < \varepsilon$ and $|a_{k_m l_p} - \beta_{l_p}| < \varepsilon$ for all $p < m$, which holds for all large k; then choose $l_m > l_{m-1}$ with $|\beta_{l_m} - L_2| < \varepsilon$ and $|a_{k_p l_m} - \alpha_{k_p}| < \varepsilon$ for all $p \le m$, which holds for all large l. For $p \le m$ the value $a_{k_p l_m}$ is then within 2ε of L_1, and for $p > m$ it is within 2ε of L_2: the preparations ρ_{k_p} and the probes σ_{l_m} form an infinite ladder at margin $|L_1 - L_2| - 4\varepsilon$, with threshold $(L_1 + L_2)/2$. □

Proposition 2.2. *Let A and B be self-adjoint operators on H_1 and on H_2 with orthonormal bases of eigenvectors. The equality test $[A = B]$ is stable. If the spectra of A and B are both unbounded above, the order test $[A < B]$ is unstable.*

Proof. For the equality test, the expectation is $\varphi_{[A=B]}(\rho, \sigma) = \sum_\lambda \operatorname{Tr}(\rho \Pi^A_\lambda) \operatorname{Tr}(\sigma \Pi^B_\lambda) = \langle u(\rho), v(\sigma) \rangle$, where λ ranges over the eigenvalues of A, and $u(\rho) = (\operatorname{Tr}(\rho \Pi^A_\lambda))_\lambda$ and $v(\sigma) = (\operatorname{Tr}(\sigma \Pi^B_\lambda))_\lambda$ are vectors of norm at most 1 in the space ℓ^2 of square-summable families indexed by these eigenvalues. Let $(\rho_k)_{k \in \mathbb{N}}$ and $(\sigma_l)_{l \in \mathbb{N}}$ be sequences of Alice's and of Bob's density operators along which both iterated limits of $\varphi_{[A=B]}(\rho_k, \sigma_l)$ exist. Bounded sequences in a Hilbert space have weakly convergent subsequences, and passing to subsequences does not change iterated limits that exist; we may therefore assume that $u(\rho_k) \to \xi$ weakly as $k \to \infty$ and $v(\sigma_l) \to \xi'$ weakly as $l \to \infty$, and then both iterated limits equal $\langle \xi, \xi' \rangle$. For the order test, the spectra of A and B, being unbounded above, allow us to choose eigenvalues $\lambda_1 < \lambda'_1 < \lambda_2 < \lambda'_2 < \cdots$, with λ_k an eigenvalue of A and λ'_l one of B, and unit eigenvectors ψ_k of A for λ_k and χ_l of B for λ'_l, for all $k, l \ge 1$. By this choice, the inequality $\lambda_k < \lambda'_l$ holds exactly when $k \le l$, and the value $\varphi_{[A<B]}(\psi_k, \chi_l)$ is therefore 1 for $k \le l$ and 0 for $k > l$; the vectors ψ_k and χ_l thus form an infinite ladder at margin 1 with threshold $\frac{1}{2}$, with the two inequalities of the definition exchanged, and the order test is unstable, by Theorem 2.1. □

The instability of $[A < B]$ is the classical failure of the double limit condition for the order of $\mathbb{N}$, carried into a Hilbert space by two orthonormal sequences; it owes nothing to noncommutativity. The equality test is stable for the opposite reason: the expectation of the equality test is an inner product, and the inner product has no infinite ladders.

The model-theoretic reading. In the theory of the structure $\mathcal{M}_C$ of Variatio 1, the order property of the expectation formula φ is the existence of ladders of C of every finite length at a fixed margin, since an infinite ladder at margin δ in a model of the theory yields, by elementary equivalence with $\mathcal{M}_C$, ladders of C of every finite length at every smaller margin. Hence φ is stable in the theory of $\mathcal{M}_C$ exactly when C is *uniformly stable*, that is, when, at every margin, the ladders of C have bounded length (see Variatio 7). Stability of C asks less, since by Theorem 2.1 it forbids infinite ladders in $\mathcal{M}_C$ alone, and in Example 7.2 we show that it is strictly weaker.

For the physicist or operator theorist. For a single bounded function, Shelah's order property is the existence of finite staircases of every length at a fixed gap, in the form of Krivine and Maurey [**KM81**], which Iovino carried into real-valued model theory [**Iov99a**]; a logician calls the expectation stable *in the theory* of the structure $\mathcal{M}_C$ when no ultrapower of $\mathcal{M}_C$ carries an infinite staircase.

For the logician. A unit vector in the range of the spectral projection Π_λ of a Hamiltonian describes a state of definite energy λ: in this state a measurement of energy returns λ with certainty. The tests $[A = B]$ and $[A < B]$ are yes-no measurements on the pair of systems, and, for each of them, the value $\varphi_C(\rho, \sigma)$, with C the test, is the probability of the answer yes when Alice's system is prepared in ρ and Bob's in σ. For the equality test this probability is an inner product of two vectors of ℓ^2, namely $\sum_\lambda p_\lambda q_\lambda$, where p_λ and q_λ are the probabilities that Alice's system and Bob's system, respectively, have energy λ; for the order test, eigenstates of increasing energy, taken alternately from Alice's system and from Bob's, produce answers that are certain: yes when Alice's eigenstate comes before Bob's in the alternation, and no otherwise. Both tests commute with every measurement of energy on either system.

A ladder is the combinatorial configuration behind a failure of the double limit condition: two disagreeing iterated limits yield an infinite ladder, and Grothendieck's double limit condition is exactly the absence of infinite ladders.

VARIATIO 3

a 1 Clav. Canone all'Unisuono

Formulas and observables

The ladders of a bipartite observable are, by definition, the ladders of a formula in the structure that the observable defines, and the correspondence does not stop at configurations: types, elementary extensions and definitions of types have physical counterparts as well. In this variation we set out this dictionary of stability (see Table 2). Each theorem is stated once and read twice, by the logician and by the physicist, and the same words carry a different meaning in each reading.

The logician's reading. Let $\varphi(x, y)$ be a formula, let $\mathcal{M}$ be a structure, and let p be a type in x over $\mathcal{M}$ that is the limit, along an ultrafilter $\mathcal{W}$ on a set I, of the types of elements a_i of $\mathcal{M}$, for $i \in I$; we write $\varphi(p, b) = \lim_{i,\mathcal{W}} \varphi^{\mathcal{M}}(a_i, b)$ for the value that p assigns to the formula $\varphi(x, b)$. The φ-type of p is *definable* if the function $b \mapsto \varphi(p, b)$ on $\mathcal{M}$ extends to a continuous function on the space $S_y(\mathcal{M})$ of types in y over $\mathcal{M}$, which is then called its *definition*. If the formula has no infinite ladder in $\mathcal{M}$, then every φ-type over $\mathcal{M}$ is definable, and the formula is stable in the theory of $\mathcal{M}$ exactly when every φ-type over every model of the theory is definable [**BY14**].

In the notation of Variatio 1, let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence of Alice's density operators, let $a = [\rho_n]_{\mathcal{U}}$ be its class in the ultrapower $\mathcal{M}_C^{\mathcal{U}}$ of the structure $\mathcal{M}_C$, and write $\mathrm{tp}(a/\mathcal{M}_C)$ for the type of a over $\mathcal{M}_C$. Let b be an element of Bob's sort in an elementary extension $\mathcal{N}$ of $\mathcal{M}_C$, and let a' be the class of $(\rho_n)_{n\in\mathbb{N}}$ in $\mathcal{N}^{\mathcal{U}}$; its type over $\mathcal{N}$ is the coheir of $\mathrm{tp}(a/\mathcal{M}_C)$ given by the limit along $\mathcal{U}$ of the types of the elements ρ_n. The value $\varphi(a', b)$, which is the value with Bob's system idealized first when b is the class of a sequence of Bob's density operators, depends only on the type of b over $\mathcal{M}_C$; this function of the type is continuous exactly when the φ-type of a is definable and its definition, evaluated at the type of b, gives $\varphi(a', b)$ for every such b, that is, when the coheir follows the definition. Definability alone is not enough: for the energy comparison, as we show in the paragraph "Both readings" below, the φ-type of the class of $(|k\rangle_A)_{k\in\mathbb{N}}$ is definable, but its coheir does not follow its definition.

The physicist's reading. Let C be a bipartite observable on the tensor product $H_1 \otimes H_2$ of the Hilbert spaces of Alice's and Bob's systems, and let ω_1 be a state of Alice's algebra $B(H_1)$, the weak* limit of a net $(\rho_i)_{i\in\mathcal{I}}$ of her density operators. Its response on actual probes, the function $\sigma \mapsto \lim_i \varphi_C(\rho_i, \sigma)$ on Bob's density operators, is given by an operator: as we saw in Variatio 1, there is a unique bounded operator Z_{ω_1} on H_2 with $\lim_i \varphi_C(\rho_i, \sigma) = \mathrm{Tr}(\sigma Z_{\omega_1})$ for every density operator σ on H_2. This response is *represented by an observable* if the same operator also gives the value at idealized probes, in the order with Bob's system idealized first: we have $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \omega_2(Z_{\omega_1})$ for every state ω_2 of $B(H_2)$, where $\otimes_{\mathsf{B}}$ is the product of Variatio 1 with Bob's system idealized first. In Theorem 3.1 below, in which $\otimes_{\mathsf{A}}$ is the product of Variatio 1 with Alice's system idealized first, we show that the observable C is stable exactly when every idealized response of C is represented by an observable.

The logician	The physicist
a structure with two sorts	the preparation spaces of Alice's and Bob's systems
an element of a sort	an actual preparation, a density operator
a formula $\varphi(x, y)$	the expectation $\varphi_C(\rho, \sigma)$ of an observable
an elementary extension, such as an ultrapower	an enlargement containing idealized preparations, classes of sequences of actual ones
the φ-type over $\mathcal{M}_C$ of an element of an ultrapower	the response on actual probes of a state of $B(H_1)$, obtained as an ultrafilter limit of a sequence of preparations
the φ-type of an element a of the structure, a function of the elements b of the other sort	the response of a preparation to the probes
a definable φ-type whose coheirs follow its definition	a Bob-first idealized response represented on all probes by one observable
realizing one coheir before the other	idealizing one system before the other
an independent amalgam, read at φ	a candidate product value at the observable C
the order property in the theory	ladders of every finite length at some fixed positive margin
a formula with no infinite ladder in the given model	a stable observable: its two orders of idealization agree
a formula stable in the theory of the model	a uniformly stable observable: bounded ladder lengths at every positive margin

TABLE 2. The dictionary of stability, read for the expectation predicate φ in the preparation structure $\mathcal{M}_C$.

Theorem 3.1 (Grothendieck [Gro52]). *For a bipartite observable C on $H_1 \otimes H_2$, the following are equivalent:*

(1) *the observable C is stable;*
(2) *the two orders of idealization agree: for every state ω_1 of $B(H_1)$ and every state ω_2 of $B(H_2)$, we have $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = (\omega_1 \otimes_{\mathsf{B}} \omega_2)(C)$;*
(3) *every idealized response of C is represented by an observable.*

PROOF. One argument serves both readings. As we saw in Variatio 1, the product with Alice's system idealized first is always $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = \omega_2(Z_{\omega_1})$; conditions (2) and (3) therefore say the same thing. If (2) fails at a pair of states ω_1 and ω_2, take nets $(\rho_i)_{i \in \mathcal{I}}$ and $(\sigma_j)_{j \in \mathcal{J}}$ of density operators converging to ω_1 and to ω_2, respectively; the iterated limits $L_1 = \lim_i \lim_j \varphi_C(\rho_i, \sigma_j)$ and $L_2 = \lim_j \lim_i \varphi_C(\rho_i, \sigma_j)$ are the two products at C, and they differ. Arguing as in the proof of Theorem 2.1, with "for all large k" and "for all large l"

replaced by "eventually along the net", we choose indices alternately: each step imposes finitely many conditions, all of which hold eventually along a net, and the choice yields sequences $(\rho_{i_p})_{p\geq 1}$ and $(\sigma_{j_m})_{m\geq 1}$ that form an infinite ladder; by Theorem 2.1, the observable C is unstable. Conversely, if C is unstable, there are sequences $(\rho_k)_{k\in\mathbb{N}}$ of Alice's density operators and $(\sigma_l)_{l\in\mathbb{N}}$ of Bob's whose iterated limits $L_1 = \lim_k \lim_l \varphi_C(\rho_k, \sigma_l)$ and $L_2 = \lim_l \lim_k \varphi_C(\rho_k, \sigma_l)$ exist and differ; their limits ω_1 and ω_2 along a nonprincipal ultrafilter on $\mathbb{N}$ are limits of subnets, along which ordinary limits persist, and therefore $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = L_1$ and $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = L_2$. □

The logician's reading. Whenever the φ-type over $\mathcal{M}_C$ determined by ω_1 is definable, its definition is given by the operator Z_{ω_1}: for a nonprincipal ultrafilter $\mathcal{V}$ on $\mathbb{N}$ and a sequence $(\sigma_m)_{m\in\mathbb{N}}$ of Bob's density operators, the definition sends the type over $\mathcal{M}_C$ of the element $[\sigma_m]_{\mathcal{V}}$ of $\mathcal{M}_C^{\mathcal{V}}$ to $\omega_2(Z_{\omega_1})$, where $\omega_2 = \lim_{m,\mathcal{V}} \sigma_m$. Ben Yaacov observed that Grothendieck's theorem of 1952, which we stated in the Aria, yields the definability of types in stable theories once it is read for formulas [**BY14**].

The physicist's reading. For the energy comparison P of Variatio 1, the projection that asks whether Bob's oscillator has the higher energy, let $|0\rangle_A, |1\rangle_A, \dots$ and $|0\rangle_B, |1\rangle_B, \dots$ be the energy eigenvectors of Alice's and of Bob's oscillator, let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$, and take $\omega_1 = \lim_{k,\mathcal{U}} |k\rangle_A$. The effective observables $T_P(|k\rangle_A) = \sum_{l>k} |l\rangle_B\langle l|_B$, by which P acts on Bob's system once Alice has prepared $|k\rangle_A$, decrease to 0 strongly, that is, the vector $T_P(|k\rangle_A)\xi$ tends to 0 for every $\xi \in H_2$, although each of these operators has norm 1; hence the value $\mathrm{Tr}(\sigma\, T_P(|k\rangle_A))$ tends to 0 for every density operator σ, and $Z_{\omega_1} = 0$: on every actual probe, the response of ω_1 takes the value 0. At the idealized probe $\omega_2 = \lim_{l,\mathcal{U}} |l\rangle_B$, however, the order with Bob's system idealized first gives $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(P) = 1$, while $\omega_2(Z_{\omega_1}) = 0$. No operator reproduces this response.

Both readings. By Theorem 3.1, Grothendieck's condition, that the iterated limits agree, holds exactly when every idealized response is represented by an observable; in the logician's reading, it holds exactly when, for every nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$ and every element a of Alice's sort in $\mathcal{M}_C^{\mathcal{U}}$, the φ-type of a is definable and the coheir of the type of a follows its definition. For the energy comparison P, the φ-type over $\mathcal{M}_P$ of the class of $(|k\rangle_A)_{k\in\mathbb{N}}$, which describes Alice's oscillator excited to infinite energy, is definable, with the constant 0 as its definition, but its coheir takes the value 1 at the class of Bob's sequence $(|l\rangle_B)_{l\in\mathbb{N}}$; accordingly, the response of $\lim_{k,\mathcal{U}} |k\rangle_A$ is not represented by an observable.

VARIATIO 4

a 1 Clav.

Hilbert space

Two experiments, one with an auxiliary qubit and one with a beam splitter, compare two unknown preparations of the same system, and both measure their overlap, the Hilbert–Schmidt inner product of the two density operators. In this variation we show that every observable whose expectation is an inner product is stable, and we recall that the separable infinite-dimensional Hilbert space is the Fraïssé limit of the finite-dimensional Hilbert spaces.

To describe the two experiments, recall that a *qubit* is a quantum system described by the Hilbert space $\mathbb{C}^2$; physicists write $|0\rangle$ and $|1\rangle$ for its standard orthonormal basis, and $|\psi\rangle$ for a vector ψ. In the *swap test*, two *registers*, quantum systems described by the same Hilbert space K, are prepared in the states given by density operators ρ and σ on K, respectively, and are coupled to an auxiliary qubit, the *ancilla*, prepared in $|+\rangle$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. A unitary operator, that is, an invertible linear isometry, then acts on the registers and the ancilla together: it exchanges the two registers on the subspace in which the ancilla is in $|1\rangle$, and it acts as the identity on the subspace in which the ancilla is in $|0\rangle$; finally, the ancilla is measured in the basis $|+\rangle$, $|-\rangle$, and the outcome $|+\rangle$ counts as acceptance. The test accepts with probability $\frac{1}{2}(1 + \operatorname{Tr}(\rho\sigma))$ [**BCWdW01**]. In optics the quantity $\operatorname{Tr}(\rho\sigma)$ is measured without an ancilla. A photon is described by $\mathbb{C}^2 \otimes K$, where the factor $\mathbb{C}^2$ records which of two paths, or *ports*, it occupies and K now describes its internal degrees of freedom, such as polarization or spatial mode; an ideal lossless balanced beam splitter acts on the port factor, leaving the internal factor unchanged, by the unitary operator $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$, and the joint state of two photons lies in the symmetric subspace of $(\mathbb{C}^2 \otimes K)^{\otimes 2}$, because photons are identical bosons, particles whose joint states are unchanged when two of them are exchanged. Two independently prepared single photons whose internal states are given by ρ and σ enter such a beam splitter, one in each input port. The internal states include all degrees of freedom relevant to distinguishability. With ideal output counting that does not resolve these degrees of freedom, they leave by different ports, an event called a *coincidence*, with probability $\frac{1}{2}(1 - \operatorname{Tr}(\rho\sigma))$, which vanishes for equal pure states. This is the Hong–Ou–Mandel effect [**HOM87**]. Both experiments thus measure the overlap $\operatorname{Tr}(\rho\sigma)$, the Hilbert–Schmidt inner product of ρ and σ.

Lemma 4.1 (The swap trick). *Let K be a Hilbert space, and let F be the unitary operator on $K \otimes K$ with $F(\psi \otimes \chi) = \chi \otimes \psi$ for all $\psi, \chi \in K$. Then, for all density operators ρ and σ on K, we have $\operatorname{Tr}((\rho \otimes \sigma)F) = \operatorname{Tr}(\rho\sigma)$. The probability of acceptance in the swap test and the probability of a coincidence in the Hong–Ou–Mandel experiment are, respectively, the expectations at (ρ, σ) of the orthogonal projections $\frac{1}{2}(I + F)$ and $\frac{1}{2}(I - F)$.*

Proof. For $\rho = |\psi\rangle\langle\psi'|$ and $\sigma = |\chi\rangle\langle\chi'|$, the operator $(\rho \otimes \sigma)F$ is $|\psi \otimes \chi\rangle\langle\chi' \otimes \psi'|$, whose trace is $\langle\chi', \psi\rangle\langle\psi', \chi\rangle = \operatorname{Tr}(\rho\sigma)$; the general case follows by linearity and continuity. □

Theorem 4.2. *Let C be a bipartite observable. If there are bounded maps u and v, from the sets $\mathfrak{D}(H_1)$ and $\mathfrak{D}(H_2)$ of density operators of Alice and of Bob into a Hilbert space, with $\varphi_C(\rho,\sigma) = \langle u(\rho), v(\sigma)\rangle$ for all $\rho \in \mathfrak{D}(H_1)$ and $\sigma \in \mathfrak{D}(H_2)$, then the observable C is stable.*

Proof. Let $(\rho_k)_{k\in\mathbb{N}}$ and $(\sigma_l)_{l\in\mathbb{N}}$ be sequences of Alice's and of Bob's density operators along which both iterated limits of $\varphi_C(\rho_k,\sigma_l)$ exist. Arguing as in the proof of Proposition 2.2, we may assume that $u(\rho_k) \to \xi$ weakly as $k \to \infty$ and $v(\sigma_l) \to \xi'$ weakly as $l \to \infty$. For each k, the inner limit over l is then $\langle u(\rho_k), \xi'\rangle$, which tends to $\langle \xi, \xi'\rangle$; by symmetry, the other order yields the same value. □

By Lemma 4.1, the expectation of the swap F is the Hilbert–Schmidt inner product $\operatorname{Tr}(\rho\sigma)$, and the swap is therefore stable, by Theorem 4.2. The expectations of the two projections $\frac{1}{2}(I \pm F)$ are $\frac{1}{2} \pm \frac{1}{2}\operatorname{Tr}(\rho\sigma)$, and the projections are stable as well, because iterated limits commute with affine functions.

A natural guess is that the stability of the swap comes from approximation by sums of products $A \otimes B$ of observables, one on each system, which the two laboratories can measure separately and combine afterwards. The norm closure of the finite sums of such products is the minimal tensor product $B(H_1) \otimes_{\min} B(H_2)$, and the distance of an observable from this closure bounds the failure of stability.

Theorem 4.3. *Let C be a bipartite observable, and let d be the distance, in the operator norm, from C to the finite sums of products $A \otimes B$ with $A \in B(H_1)$ and $B \in B(H_2)$.*

(a) *For all sequences $(\rho_k)_{k\in\mathbb{N}}$ of Alice's density operators and $(\sigma_l)_{l\in\mathbb{N}}$ of Bob's along which both iterated limits of $\varphi_C(\rho_k,\sigma_l)$ exist, the two limits differ by at most $2d$.*

(b) *Every self-adjoint element of $B(H_1) \otimes_{\min} B(H_2)$ is stable; in particular, every compact bipartite observable is stable, and, if H_1 or H_2 is finite-dimensional, every bipartite observable is stable.*

(c) *If $C \in B(H_1) \otimes_{\min} B(H_2)$, then the effective map $\rho \mapsto T_C(\rho)$, extended linearly to the trace-class operators on H_1 with the trace norm, is a compact operator into $B(H_2)$.*

Proof. (a) For a finite sum $C' = \sum_{j=1}^{r} A_j \otimes B_j$ of products, the expectation $\varphi_{C'}(\rho,\sigma) = \sum_j \operatorname{Tr}(\rho A_j)\operatorname{Tr}(\sigma B_j)$ is a sum of products of a function of ρ and a function of σ; along subsequences on which every factor converges, both iterated limits of $\varphi_{C'}$ equal the same sum of products of limits. Given sequences along which both iterated limits of φ_C exist, pass to such subsequences; since $|\varphi_C - \varphi_{C'}| \le \|C - C'\|$, each iterated limit of φ_C lies within $\|C - C'\|$ of the common value for C', and the two differ by at most $2\|C - C'\|$. Taking the infimum over C' gives (a).

(b) The distance d vanishes on the norm closure. The rank-one operators $|\psi \otimes \chi\rangle\langle\psi' \otimes \chi'| = |\psi\rangle\langle\psi'| \otimes |\chi\rangle\langle\chi'|$ span a dense subspace of the compact operators, and, if H_1 has an orthonormal basis $e_1, \dots, e_n$, then $C = \sum_{i,j\le n} |e_i\rangle\langle e_j| \otimes C_{ij}$ for suitable $C_{ij} \in B(H_2)$; the case of a finite-dimensional H_2 is symmetric.

(c) For a product, $T_{A\otimes B}(\tau) = \operatorname{Tr}(\tau A)\, B$ has rank at most one, and the norm of $T_C - T_{C'}$ as an operator from the trace-class operators into $B(H_2)$ is at most $\|C - C'\|$; norm limits of operators of finite rank are compact. □

If K is infinite-dimensional, the swap F lies outside the minimal tensor product: the effective map of F is $T_F(\rho) = \rho$, which sends the pure states of an orthonormal sequence to operators at mutual distance 1, a sequence with no convergent subsequence, and the conclusion of part (c) of Theorem 4.3 fails for F. The stability of the swap comes instead from Theorem 4.2, since its expectation is an inner product.

The Fraïssé limit

Fraïssé's theorem [Fra54] gives a countable homogeneous limit for every class of finitely generated structures with countably many isomorphism types that is closed under substructures and has the joint embedding and amalgamation properties. Ben Yaacov extended the theorem to metric structures [BY15], under metric forms of these conditions, and the limit is then a separable complete structure that is approximately homogeneous. The finite-dimensional Hilbert spaces form a class of this kind in the metric sense: the class is closed under subspaces, and two isometric embeddings of a space E into spaces E_1 and E_2 amalgamate in the space $E_1 \oplus_E E_2$, in which the orthogonal complements of E in E_1 and in E_2 are orthogonal to each other; joint embedding is the case $E = \{0\}$. The limit is the separable infinite-dimensional Hilbert space, which is homogeneous in the exact sense: every isometry between finite-dimensional subspaces extends to a unitary operator on the whole space.

The model-theoretic reading. Let H be the separable infinite-dimensional Hilbert space, and let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$. For the formula $\mathrm{Re}\langle x, y\rangle$ and a bounded sequence $(\xi_n)_{n\in\mathbb{N}}$ in H, the restriction to this formula of the type over H of the class $[\xi_n]_{\mathcal{U}}$ in the ultrapower $H^{\mathcal{U}}$ is the function $y \mapsto \mathrm{Re}\langle \xi, y\rangle$, where ξ is the weak limit of $(\xi_n)_{n\in\mathbb{N}}$ along $\mathcal{U}$; this function is given by the element ξ of H, and the type is therefore definable (see Variatio 3). The expectation of an overlap measurement in an ultrapower is again an inner product, of vectors of an ultrapower of a Hilbert space, and the argument of Theorem 4.2 shows that it has no infinite ladder; by Łoś's theorem, the ladders of overlap measurements therefore have bounded length at every margin, and in Variatio 8 we give an explicit bound. Over a closed subspace E of an elementary extension of H, vectors a and b of that extension are independent in the sense of nonforking (see Variatio 6) exactly when $a - P_E a$ and $b - P_E b$ are orthogonal, where P_E is the orthogonal projection onto E [BYBHU08]; in the amalgam $E_1 \oplus_E E_2$, the spaces E_1 and E_2 are independent over E.

For the physicist or operator theorist. The type of a vector over H, restricted to the inner product, is the weak limit of the sequence that represents the vector, and independence over a subspace E is orthogonality of the components orthogonal to E.

For the logician. The swap test and the Hong–Ou–Mandel experiment both perform the two-outcome measurement given by the orthogonal projections $\frac{1}{2}(I + F)$ and $\frac{1}{2}(I - F)$ of Lemma 4.1. For pure states ψ and χ, the overlap $\mathrm{Tr}(\rho\sigma)$ of $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\chi\rangle\langle\chi|$ is $|\langle\psi, \chi\rangle|^2$, the probability that a system prepared in ψ passes the test for χ, the measurement given by the projections $|\chi\rangle\langle\chi|$ and $I - |\chi\rangle\langle\chi|$. Since the two experiments measure the pair and never either system alone, they compare two unknown states without determining either of them.

For an inner product, the weak limits of Alice's and of Bob's vectors are again vectors of the Hilbert space, and both iterated limits equal their inner product. For inner-product responses, idealization never leaves the Hilbert space, although idealized vectors do: an orthonormal sequence has weak limit 0, while its class in the ultrapower has norm 1. This is why Hilbert space is the Fraïssé limit associated with stability.

VARIATIO 5

a 1 ô vero 2 Clav.

Stability in seven languages

The double limit condition has been found, under different names, in analysis, Banach space theory, the theory of Banach algebras, probability and model theory. For a bipartite observable, each of these forms becomes a statement about idealized preparations. The theorem below collects seven of them, and a single example is then read in all seven. Two of these forms, the agreement of the two orders of idealization and the representation of idealized responses by observables, are already in Theorem 3.1 of Variatio 3; three others describe operations of a laboratory, and each of them is developed in a later variation: the independent product in Variatio 6, the simulation of idealized preparations by finite mixtures in Variatio 9, and randomization in Variatio 10. Between the first two of these variations, in Variations 7 and 8, we treat the uniform version of the condition, which the ultrapower turns into a statement about every elementary extension.

Throughout this variation, the letter C denotes a bipartite observable on the tensor product $H_1 \otimes H_2$ of Alice's and Bob's Hilbert spaces, with expectation $\varphi_C(\rho, \sigma) = \mathrm{Tr}((\rho \otimes \sigma)C)$ and with effective observable $T_C(\rho)$ for each density operator ρ of Alice, the observable that C induces on Bob's system once Alice has prepared ρ; the symbols $\otimes_{\mathsf{A}}$ and $\otimes_{\mathsf{B}}$ denote the products of Variatio 1, given by the two orders of idealization, with Alice's system and with Bob's system idealized first, respectively.

Theorem 5.1. *For every bipartite observable C on $H_1 \otimes H_2$, the following are equivalent:*

(A1) *the observable C is stable: for all sequences $(\rho_k)_{k\in\mathbb{N}}$ of Alice's density operators and $(\sigma_l)_{l\in\mathbb{N}}$ of Bob's, the two iterated limits of $\varphi_C(\rho_k, \sigma_l)$ agree whenever both exist;*

(A2) (Grothendieck) *the effective observables $T_C(\rho)$, for ρ in the set $\mathfrak{D}(H_1)$ of Alice's density operators, form a relatively weakly compact subset of the Banach space $B(H_2)$;*

(A3) (Iovino; Davis, Figiel, Johnson, Pełczyński) *the expectation factors through the pairing of a reflexive Banach space X with its dual X^*: there are bounded affine maps u from $\mathfrak{D}(H_1)$ into X and v from $\mathfrak{D}(H_2)$ into X^* with $\varphi_C(\rho, \sigma) = v(\sigma)(u(\rho))$ for all ρ and σ;*

(A4) (Arens, Pym, Ülger) *the two orders of idealization agree on all pairs of states: the two products satisfy $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = (\omega_1 \otimes_{\mathsf{B}} \omega_2)(C)$ for every state ω_1 of $B(H_1)$ and every state ω_2 of $B(H_2)$;*

(A5) (Mazur) *every idealized preparation of Alice is simulated, to any accuracy and uniformly over all probes, idealized ones included, by finite mixtures of actual preparations, with Bob's system idealized first;*

(A6) (Pták, Simons) *for any two finitely additive probability measures, one on $\mathfrak{D}(H_1)$ and one on $\mathfrak{D}(H_2)$, the two iterated integrals of the expectation against them agree;*

(A7) (Harrington) *for every state of Alice's algebra and every state of Bob's algebra, the two orders of realizing the pair give the same value, the value at C of the independent product of the two states.*

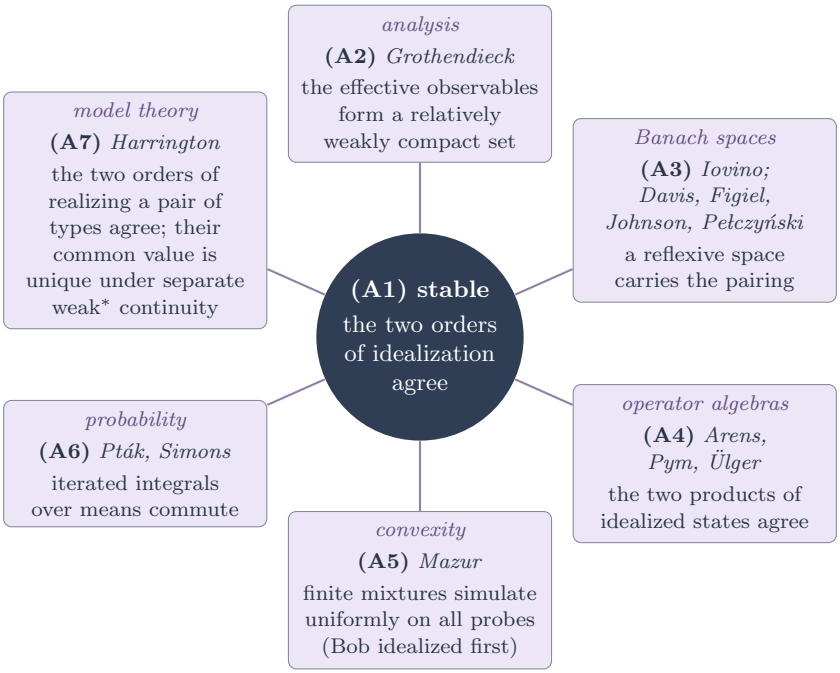


FIGURE 3. Seven readings of one condition, the forms of stability in Theorem 5.1, each coming from a different field.

The equivalence of the seven conditions

We prove Theorem 5.1 in section P.1. The proof is organized around (A4), the agreement of the two products: each of the other conditions is joined to it, or to a condition already joined to it; most of the links are classical theorems of the fields from which the conditions come. The equivalence of (A1) and (A4) is Theorem 3.1, whose proof extracts a ladder from two disagreeing iterated limits. For (A2), the product with Bob's system idealized first is given by the bitranspose of the effective map. By Gantmacher's theorem, a bounded operator between Banach spaces is weakly compact exactly when the bitranspose of the operator takes values in the original target space [**Gan40**], and this theorem turns (A4) into weak compactness. The compactness in (A2) is for the weak topology of the Banach space $B(H_2)$, which is finer than the weak operator topology, and under (A2) every idealized response is therefore represented by an observable. Condition (A3) follows from (A2): the map $\rho \mapsto T_C(\rho)$, extended linearly to the trace-class operators on H_1, is then a weakly compact operator into $B(H_2)$, and, by the Davis–Figiel–Johnson–Pełczyński theorem, it factors through a reflexive space [**DFJP74**]. Iovino showed that every bounded function with the double limit property factors, by bounded maps, through the pairing of a reflexive space with its dual [**Iov99b**]. Conversely, (A3) yields (A1), because bounded subsets of a reflexive space are relatively weakly compact. Condition (A5) rests on Mazur's theorem [**Maz33**]: a convex set has the same weak and norm closures, and a weak approximation of the operator Z_{ω_1} by effective observables of mixtures thus becomes a norm approximation, uniform over probes. In (A6), an iterated integral of a function that is affine in each variable depends only on the barycenters of the two measures, which are states, and the two iterated integrals are therefore the two products of (A4); the equivalence of the double limit condition with the agreement of iterated integrals, for arbitrary bounded functions, is the Pták–Simons theorem [**Ptá63**, **Sim68**]. Finally, for (A7), the two orders of realizing a pair of types are the two iterated ultrafilter limits, and their agreement is (A4) in the language of types; for a formula stable in the theory, it is Harrington's lemma [**HH84**]. As a function of the two states, the common value in (A7) is symmetric in the two systems and is the only extension of the expectation that is weak* continuous in each state separately (see Theorem 6.4).

One example in seven languages

Let P be the energy comparison of two oscillators, with $\varphi_P(|k\rangle_A, |l\rangle_B) = 1$ for $k < l$ and 0 otherwise, where $|0\rangle_A, |1\rangle_A, \dots$ and $|0\rangle_B, |1\rangle_B, \dots$ are the energy eigenvectors of Alice's oscillator and of Bob's, and let $\omega_1 = \lim_{k,\mathcal{U}} |k\rangle_A$ and $\omega_2 = \lim_{l,\mathcal{U}} |l\rangle_B$ for a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$.

(A1) The iterated limit taken first over Alice's index is 0, and the one taken first over Bob's is 1.

(A2) The effective observables $T_P(|k\rangle_A) = \sum_{l>k} |l\rangle_B\langle l|_B$ decrease to 0 in the strong operator topology, but the state ω_2 assigns every one of them the value 1 and their strong limit the value 0; no subsequence converges weakly in $B(H_2)$.

(A3) The comparison does not factor, by bounded maps, through the pairing of a reflexive space with its dual: for the vectors $u(|k\rangle_A)$ and $v(|l\rangle_B)$ of such a factorization, weak limits of subsequences would force the two iterated limits to agree.

(A4) The two products of ω_1 and ω_2 at P are 0 and 1.

(A5) No mixture ρ of Alice's preparations simulates ω_1 within any error less than 1: the response of ω_1 is 0 on every $|l\rangle_B$, while $\mathrm{Tr}((|l\rangle_B\langle l|_B)T_P(\rho)) = \sum_{k<l} {}_A\langle k|\rho|k\rangle_A$ tends to 1 as l grows.

(A6) Let μ be the finitely additive probability measure that assigns to a set of Alice's density operators the value 1 if the indices k for which $|k\rangle_A$ belongs to the set form a member of $\mathcal{U}$, and the value 0 otherwise, and let ν be defined in the same way from $(|l\rangle_B)_{l\in\mathbb{N}}$; one iterated integral against μ and ν is 0, and the other is 1.

(A7) The φ-types determined by ω_1 and ω_2 are realized in two orders that assign the comparison the values 0 and 1; the two states have no independent product at P.

The model-theoretic reading. In the structure $\mathcal{M}_C$ of Variatio 1, condition (A2) says that the coheirs of the φ-types over $\mathcal{M}_C$ of the elements of Alice's sort of ultrapowers of $\mathcal{M}_C$ follow their definitions: for $a = [\rho_n]_{\mathcal{U}}$ and the class $b = [\sigma_m]_{\mathcal{V}}$ of a sequence of Bob's density operators, along nonprincipal ultrafilters $\mathcal{U}$ and $\mathcal{V}$ on $\mathbb{N}$, the coheir value $\varphi(a', b)$ of Variatio 3, the value with Bob's system idealized first, is $\omega_2(Z_{\omega_1})$, where $\omega_1 = \lim_{n,\mathcal{U}} \rho_n$ and $\omega_2 = \lim_{m,\mathcal{V}} \sigma_m$, and depends continuously on the type of b over $\mathcal{M}_C$. Condition (A6) concerns Keisler measures: for a sequence $(\mu_n)_{n\in\mathbb{N}}$ of finitely supported probability measures on Alice's sort of $\mathcal{M}_C$, the functional $\theta(x, b) \mapsto \lim_{n,\mathcal{U}} \int \theta(x, b)\, d\mu_n(x)$ is a Keisler measure over $\mathcal{M}_C$, and condition (A6), read for such measures on the two sorts, says that the two products of Keisler measures agree at φ. Condition (A7) is the counterpart, for the single structure $\mathcal{M}_C$, of Harrington's lemma (see Theorem 6.1): the two independent amalgams obtained in an iterated ultrapower give φ the same value.

For the physicist or operator theorist. In (A2), the logician's *definition* of the type of an idealized preparation is the operator Z_{ω_1}, read as a continuous function of the probe, and the condition that coheirs follow definitions says that this operator also gives the values at idealized probes. A Keisler measure is a finitely additive probability measure on preparations, seen through the formulas, and (A6) is Fubini's theorem for such measures.

For the logician. The conditions that describe operations of a laboratory are (A5), (A6) and (A7). A finite mixture $\sum_{i=1}^{r} p_i \rho_i$ of preparations $\rho_1, \ldots, \rho_r$, with weights $p_i \geq 0$ summing to 1, is prepared by choosing the preparation ρ_i at random with probability p_i and forgetting the choice, and condition (A5) says that, for every $\varepsilon > 0$, such a procedure reproduces an idealized preparation within ε, simultaneously for all of Bob's probes, in the order with Bob's system idealized first; by Theorem 5.1, this happens exactly for stable observables. Condition (A6) concerns preparations drawn at random from infinitely many (see Variatio 10), and condition (A7) concerns the joint state of two systems prepared independently, which is $\rho \otimes \sigma$ for actual preparations ρ and σ and must be defined by a limit for idealized ones. For the energy comparison, all three fail with a gap of 1, the largest that a projection allows, and for one reason: the event "Bob's oscillator has the higher energy" has probability 0 or 1 according to the order in which the two oscillators, excited to infinite energy, are approached.

The seven conditions are seven readings of Grothendieck's double limit condition, and the staircase of the Aria, the matrix of the energy comparison, violates all seven for one reason: its two iterated limits are 0 and 1.

VARIATIO 6

a 1 Clav. Canone alla Seconda

The independent product

Two systems are prepared independently. If the preparations are actual, the joint state is the tensor product of the two density operators; if they are idealized, there are two candidates for the joint state, the two products of Variatio 1, obtained by placing the two systems, in either order, in an iterated ultrapower. In this variation we show that, at stable observables, the two candidates coincide and assemble into a symmetric product state, the only extension of the product of actual preparations that is weak* continuous in each marginal state separately (see Theorem 6.4); the logician reads the same exchange of limits in Harrington's lemma, for two-valued types over a model.

Let C be a bipartite observable, and recall from Variatio 1 the structure $\mathcal{M}_C$, whose two sorts are the sets of density operators of Alice and of Bob and whose predicate symbol φ stands for the expectation of C. Let p and q be types over $\mathcal{M}_C$ in Alice's sort and in Bob's, respectively, realized by elements a and b of ultrapowers of $\mathcal{M}_C$, as in Variatio 1. A joint type of (a, b) obtained in either order of the iterated ultrapower is an *independent amalgam* of p and q over $\mathcal{M}_C$, and in general the amalgam depends on the ultrafilters, not only on p and q. The uniqueness of independent amalgams separates stable theories from unstable ones: when φ is stable in the theory of $\mathcal{M}_C$, the coheir of a φ-type over a model does not depend on the ultrafilter and is the unique nonforking extension, for the notion of forking recalled below, and independence is symmetric [**HH84**]. For the single structure $\mathcal{M}_C$ the relevant condition is stability of C, which is weaker than uniform stability (see Theorem 3.1 and Theorem 7.1); at each stable observable, this weaker condition already makes the two candidate values agree. Simple theories retain a symmetric independence relation without uniqueness, and in Part III we show what it becomes in quantum error correction.

The logician's reading. Let M be a model of a stable theory, let p and q be complete types over M in variables x and y, and let $\varphi(x, y)$ be a formula. Stability provides definitions: there are formulas $d_p\varphi(y)$ and $d_q\varphi(x)$ over M such that, for every tuple b from M, the formula $\varphi(x, b)$ belongs to p exactly when $M \models d_p\varphi(b)$, and, for every tuple a from M, the formula $\varphi(a, y)$ belongs to q exactly when $M \models d_q\varphi(a)$.

Theorem 6.1 (Harrington's lemma [HH84]). *Let M be a model of a stable theory, let p and q be complete types over M in the variables x and y, and let $\varphi(x, y)$ be a formula. Then the formula $d_p\varphi(y)$ belongs to q if and only if the formula $d_q\varphi(x)$ belongs to p.*

Every type over a model is finitely satisfiable in it; hence there is an ultrafilter $\mathcal{U}$ on the tuples of M such that a formula $\theta(x, b)$ belongs to p exactly when the set of tuples a with $M \models \theta(a, b)$ belongs to $\mathcal{U}$, and there is an ultrafilter $\mathcal{V}$ that gives q in the same way. If φ is regarded as a $\{0, 1\}$-valued function on pairs of tuples, then the two sides of the lemma read:

$$\lim_{b,\mathcal{V}} \lim_{a,\mathcal{U}} \varphi(a, b) = 1 \qquad \text{and} \qquad \lim_{a,\mathcal{U}} \lim_{b,\mathcal{V}} \varphi(a, b) = 1.$$

Harrington's lemma thus asserts that two iterated ultrafilter limits may be exchanged. To state the role of the lemma, we recall forking (see Variatio 1): a formula divides over a set A when, along some sequence of its parameters indiscernible over A, its instances are k-inconsistent for some integer k; a type forks over A when it implies a finite disjunction of formulas that divide over A, and $a \mathop{\smile\!\!\!\!\!\!|}_A b$ means that $\mathrm{tp}(a/Ab)$ does not fork over A. In every theory, a coheir of a type over a model does not fork over the model; hence the two orders of the iterated ultrapower of Variatio 1 always produce nonforking amalgams, and in a stable theory they produce the only one. Harrington's lemma is the step that makes nonforking independence symmetric in a stable theory, and, together with the uniqueness of nonforking extensions over models, it yields the characterization of stable theories, due to Harnik and Harrington, as the theories with an independence relation that is stationary over models [**HH84**].

The physicist's reading. Let $\mathcal{O}$ be the set of stable observables on the tensor product $H_1 \otimes H_2$ of Alice's and Bob's Hilbert spaces, a norm-closed real vector space of self-adjoint operators that contains I (see Proposition 6.2 below). For a state ω_1 of $B(H_1)$ and a state ω_2 of $B(H_2)$, the two orders of idealization agree at every $C \in \mathcal{O}$, by Theorem 3.1, and we write $\omega_1 \otimes \omega_2$ for the functional on $\mathcal{O}$ whose value at C is their common value. We seek a state on this space of observables that extends the product of actual preparations and whose value at each $C \in \mathcal{O}$ depends weak* continuously on each marginal state separately. These requirements determine the scalar extension at each stable observable, because actual states are dense. The theorem below shows that the resulting functional is a state on $\mathcal{O}$ and is symmetric under exchange of the two systems. In this reading, uniqueness concerns the separately continuous extension on stable observables, rather than arbitrary joint states with prescribed marginals.

Proposition 6.2. *The set $\mathcal{O}$ of stable observables on $H_1 \otimes H_2$ is a real vector space that contains I and every self-adjoint element of the minimal tensor product $B(H_1) \otimes_{\min} B(H_2)$, and is closed in the operator norm. The same holds for the set of uniformly stable observables, defined in Variatio 7. In particular, if P and Q are stable orthogonal projections on $H_1 \otimes H_2$, then so is $I - P$, and so is $P + Q$ when $PQ = 0$.*

Proof. The expectation φ_C depends linearly on C, and $|\varphi_C(\rho,\sigma) - \varphi_{C'}(\rho,\sigma)| \le \|C - C'\|$ for all density operators ρ and σ. Let C and C' be stable, let a and b be real numbers, and let $(\rho_k)_{k\in\mathbb{N}}$ and $(\sigma_l)_{l\in\mathbb{N}}$ be sequences along which both iterated limits of $\varphi_{aC+bC'}$ exist. A diagonal argument yields subsequences along which both iterated limits of φ_C and of $\varphi_{C'}$ exist; passing to them does not change the iterated limits of $\varphi_{aC+bC'}$, and each of these is a times the common iterated limit of φ_C plus b times that of $\varphi_{C'}$. If C is the norm limit of stable observables C_n, then, along subsequences on which the iterated limits of φ_{C_n} exist, each iterated limit of φ_C lies within $\|C - C_n\|$ of the common iterated limit of φ_{C_n}; the two iterated limits of φ_C therefore differ by at most $2\|C - C_n\|$ for every n, and they agree. The identity is stable because $\varphi_I = 1$, and the minimal tensor product is treated in Theorem 4.3. For uniform stability, the same arguments apply to the expectation of an ultrapower of the structure $\mathcal{M}_C$, by condition (B3) of Theorem 7.1. The last statement holds because $I - P$ and $P + Q$ are linear combinations of stable observables. □

Ordered by inclusion of ranges, with $I - P$ as the orthocomplement of P, the projections on $H_1 \otimes H_2$ form an orthomodular poset, and by Proposition 6.2 the stable projections form a sub-poset closed under orthocomplements and orthogonal sums. The quantum-logical conjunction does not preserve stability: the *meet* $P_1 \wedge P_2$ of two projections, the projection onto the intersection of their ranges, can be unstable when P_1 and P_2 are uniformly stable.

Proposition 6.3. *If H_1 and H_2 are infinite-dimensional, then there are orthogonal projections P_1 and P_2 in $B(H_1) \otimes_{\min} B(H_2)$, both uniformly stable, whose meet $P_1 \wedge P_2$ is unstable.*

Proof. The idea is to hide the energy comparison of the Aria in the intersection of two subspaces that are compact perturbations of a product projection, by tilting each entry of the staircase into a fresh direction through a small angle, in opposite senses for the two subspaces. Let $(e_k)_{k\geq 1}$ and $(e'_k)_{k\geq 1}$ be sequences in H_1 whose terms together form an orthonormal set, let $(f_l)_{l\geq 1}$ and $(f'_l)_{l\geq 1}$ be such sequences in H_2, and let P be the orthogonal projection onto the closed span of the vectors $e_k \otimes f_l$ with $k < l$, which is unstable, since the vectors e_k and f_l form an infinite ladder at margin 1 (see Theorem 2.1). For $k \geq l$, let $\zeta_k = 1/k$ and

$$n_{kl} = \cos\zeta_k\, e_k \otimes f_l + \sin\zeta_k\, e'_k \otimes f'_l, \qquad n'_{kl} = \cos\zeta_k\, e_k \otimes f_l - \sin\zeta_k\, e'_k \otimes f'_l,$$

and let P_1 and P_2 be the orthogonal projections onto the closed subspaces spanned by the range of P together with the vectors n_{kl}, and with the vectors n'_{kl}, respectively, for $k \geq l$. The planes spanned by $e_k \otimes f_l$ and $e'_k \otimes f'_l$, for $k \geq l$, are mutually orthogonal and orthogonal to the range of P, and in each of them the lines through n_{kl} and n'_{kl} meet only in 0, since $0 < \zeta_k < \pi/2$. Both subspaces decompose along these planes and the range of P, so their intersection is the range of P, and $P_1 \wedge P_2 = P$.

Let Q_1 and Q_2 be the orthogonal projections onto the closed spans of the vectors e_k and of the vectors f_l. The operator $P_1 - Q_1 \otimes Q_2$ is the orthogonal sum over the pairs $k \geq l$ of the rank-two operators $|n_{kl}\rangle\langle n_{kl}| - |e_k \otimes f_l\rangle\langle e_k \otimes f_l|$, of norm $\sin\zeta_k$. For every $\varepsilon > 0$, only finitely many pairs $k \geq l$ satisfy $1/k \geq \varepsilon$, so this sum is compact. Since $Q_1 \otimes Q_2$ and the compact operators lie in $B(H_1) \otimes_{\min} B(H_2)$, so does P_1, and likewise P_2; both are uniformly stable by Proposition P.3. □

By von Neumann's alternating projection theorem, the meet is the strong limit of the powers $(P_1P_2)^m$, and the convergence is in norm exactly when the sum of the two ranges is closed [**Deu01**]; in Proposition 6.3 the angles ζ_k accumulate at 0, and the sum of the ranges is not closed. The self-adjoint operators $(P_1P_2P_1)^m = (P_1P_2)^m P_1$ also converge strongly to $P_1 \wedge P_2$; they lie in the minimal tensor product, an algebra, and are therefore stable, whereas their strong limit $P_1 \wedge P_2$ is not: stability passes to norm limits, by Proposition 6.2, but not to strong limits.

Theorem 6.4. *The following hold:*

(a) (Existence) *for all states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$, the functional $\omega_1 \otimes \omega_2$ is a state of $\mathcal{O}$: linear, positive, with value 1 at I;*

(b) (Extension) *if ω_1 and ω_2 are given by density operators ρ and σ, then we have $(\omega_1 \otimes \omega_2)(C) = \mathrm{Tr}((\rho \otimes \sigma)C)$ for every $C \in \mathcal{O}$;*

(c) (Uniqueness) *for each $C \in \mathcal{O}$, the function $(\omega_1, \omega_2) \mapsto (\omega_1 \otimes \omega_2)(C)$ is the only function on pairs of states that is weak* continuous in each variable separately and satisfies* (b)*;*

(d) (Symmetry) *if $F_{12}\colon H_1 \otimes H_2 \to H_2 \otimes H_1$ is the unitary operator with $F_{12}(\xi \otimes \eta) = \eta \otimes \xi$, then, for all states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$ and every $C \in \mathcal{O}$, the observable $F_{12}CF_{12}^*$ on $H_2 \otimes H_1$ is stable and the value $(\omega_2 \otimes \omega_1)(F_{12}CF_{12}^*)$ equals $(\omega_1 \otimes \omega_2)(C)$, where the product on the left is taken on the stable observables on $H_2 \otimes H_1$.*

Proof. For states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$, the map $C \mapsto (\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = \omega_2(Z_{\omega_1})$ is linear on the bounded self-adjoint operators, because the operator Z_{ω_1} depends linearly on C; on $\mathcal{O}$ the map is $\omega_1 \otimes \omega_2$.

(a) If $C \in \mathcal{O}$ is positive, then $\varphi_C(\rho, \sigma) \geq 0$ for all density operators ρ and σ, and the iterated limit $(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C)$ of such values is nonnegative; at $C = I$ every expectation, and hence each product, equals 1.

(b) When one of the two states is actual, the two products agree, and both equal the expectation in the product state (see Variatio 1).

(c) For $C \in \mathcal{O}$, the value $(\omega_1 \otimes \omega_2)(C)$ equals $\omega_2(Z_{\omega_1})$, which is weak* continuous in ω_2, and equals $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C)$, which is weak* continuous in ω_1 by the symmetric argument; the function is therefore separately weak* continuous. Let g be another separately weak* continuous function on pairs of states that satisfies (b). For a density operator σ on H_2, the functions $g(\cdot, \sigma)$ and $(\cdot \otimes \sigma)(C)$ are weak* continuous and agree on the actual states, which are dense (see Variatio 1); they therefore agree on all states of $B(H_1)$. For a fixed state ω_1, the functions $g(\omega_1, \cdot)$ and $(\omega_1 \otimes \cdot)(C)$ are then weak* continuous and agree on the actual states of $B(H_2)$, and hence everywhere.

(d) For density operators ρ on H_1 and σ on H_2, we have $\mathrm{Tr}((\sigma \otimes \rho)F_{12}CF_{12}^*) = \mathrm{Tr}(F_{12}^*(\sigma \otimes \rho)F_{12}C) = \mathrm{Tr}((\rho \otimes \sigma)C)$, so the expectation of $F_{12}CF_{12}^*$ is that of C with the two variables exchanged. The double limit condition is symmetric in the two variables, and $F_{12}CF_{12}^*$ is therefore stable; exchanging the variables also exchanges the two orders of idealization, so $(\omega_2 \otimes_{\mathsf{A}} \omega_1)(F_{12}CF_{12}^*) = (\omega_1 \otimes_{\mathsf{B}} \omega_2)(C)$, and both sides equal the common values of the two orders. □

By Theorem 3.1, the set $\mathcal{O}$ consists exactly of the bipartite observables to which the two orders of idealization assign the same value at every pair of states.

Both readings. A type over a model is *stationary* if it has a unique nonforking extension to every larger set; in the iterated ultrapower, stationarity ensures that the type of the element realized second does not depend on the ultrafilter. In the language of independent amalgams, (a) corresponds to the existence of an independent amalgam, (c) to stationarity over a model, and (d) to the symmetry of independence. For the energy comparison, at the states that describe Alice's and Bob's oscillators excited to infinite energy, the two orders of realization yield the values 0 and 1, and uniqueness and symmetry are lost together.

On stable observables, the two orders of idealization have a common value, and these values form a symmetric product of the marginal states. For each stable observable, separate weak* continuity makes this extension unique; the double limit condition is what makes the two readings agree.

VARIATIO 7

a 1 ô vero 2 Clav. Al tempo di Giga

Discrimination and learning

Alice prepares one of many known states, and Bob distinguishes it by adaptive queries of expectation responses, each probe chosen in view of the previous answers. In the online model of learning, a learner is shown one probe at a time, predicts whether the expectation for the unknown preparation lies above or below a threshold, and is then told the correct answer. The threshold is fixed within each protocol, a positive margin separates the counted responses from it, and the bounds are uniform in the threshold. In this variation we show that, for uniformly stable observables, the number of preparations distinguished by a given number of such queries grows at most polynomially, and the learner makes boundedly many mistakes.

Trees

A ladder is a linear configuration, whereas an adaptive protocol branches, each probe depending on the outcomes of the previous ones; the configuration that records such a protocol is a tree. Let C be a bipartite observable, and fix a margin $\delta > 0$ and a threshold t. For $n \geq 1$, a *tree of height* n assigns a probe σ_s to every node s of the complete binary tree of height n and a preparation ρ_η to every branch η, in such a way that, whenever η passes through s,

$$\varphi_C(\rho_\eta, \sigma_s) \geq t + \tfrac{\delta}{2} \ \text{ if } \eta \text{ turns right at } s, \qquad \varphi_C(\rho_\eta, \sigma_s) \leq t - \tfrac{\delta}{2} \ \text{ if it turns left,}$$

or the same holds, for all branches and nodes, with the two inequalities exchanged. Thus, with n probes, each chosen in view of the previous outcomes, Bob tells apart the 2^n preparations of a tree of height n. The pairs (ρ_i, σ_i) of a ladder are its *rungs*, and, for every $h \geq 1$, a ladder with 2^h rungs contains a tree of height h: the first probe is that of the middle rung, and binary search continues in the half of the ladder consistent with the outcome. The observable C is *uniformly stable* if, at every margin, its ladders have bounded length.

Theorem 7.1. *For every bipartite observable C on $H_1 \otimes H_2$, the following are equivalent:*

(B1) *the observable C is uniformly stable;*

(B2) (Shelah) *at every margin, the trees of C have heights bounded independently of the threshold;*

(B3) *for every (equivalently, for some) nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the expectation remains stable on the ultrapower along $\mathcal{U}$;*

(B4) (Sauer, Shelah) *at every margin δ, the number of preparations that n adaptive probes tell apart is at most a polynomial in n, whose degree is the bound on the height of trees;*

(B5) (Littlestone) *at every margin, the side of a fixed threshold on which the response of an unknown preparation lies can be predicted online with a number of mistakes bounded independently of the threshold; the same holds, with a larger bound, when the threshold varies from probe to probe.*

If the conditions fail, then at some margin, for every n, there are 2^n preparations that n probes tell apart.

Two arguments

We prove Theorem 7.1 in section P.1; the proof rests on two arguments, which we describe here. The counting bound of (B4) rests on an observation about the largest height $d(\mathcal{R})$ of a tree whose preparations lie in a set $\mathcal{R}$: for any probe, the preparations in $\mathcal{R}$ that answer high and those that answer low form two disjoint parts of $\mathcal{R}$, which need not exhaust it, and the two parts cannot both carry trees of height $d(\mathcal{R})$, because two such trees, placed above the probe, would form a tree of height $d(\mathcal{R}) + 1$. By induction on n, with Pascal's rule, the number of preparations that n probes tell apart is at most $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}$, a polynomial of degree d, where $d = d(\mathcal{R})$. This is the tree form of the Sauer–Shelah lemma [**Sau72**, **She72**].

Littlestone's learner [**Lit88**] rests on the same observation: it maintains the set V of preparations consistent with the answers so far and predicts the side whose consistent preparations carry the higher tree. Each mistake decreases $d(V)$, so there are at most d mistakes. Conversely, along a tree of height n, choosing at each node the branch opposite to the prediction produces n mistakes.

For (B3), the idea is a count: trees of every height glue, in an ultrapower, into a tree whose $2^{\aleph_0}$ branches are told apart by countably many probes, whereas, if the expectation of the ultrapower were stable, the responses of the preparations at the branches to these probes would lie in a separable space. Fix a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$. Sequences $(\rho_k)_{k\in\mathbb{N}}$ of Alice's preparations and $(\sigma_k)_{k\in\mathbb{N}}$ of Bob's probes, which are families of experiments indexed by k, define elements $[\rho_k]_{\mathcal{U}}$ and $[\sigma_k]_{\mathcal{U}}$ of the ultrapower $\mathcal{M}_C^{\mathcal{U}}$ of the structure $\mathcal{M}_C$ of Variatio 1. At this pair, the expectation of the ultrapower takes the value $\lim_{k,\mathcal{U}} \varphi_C(\rho_k, \sigma_k)$. If C has trees of every height at a fixed margin, they glue along the ultrafilter into such an infinite tree. If the expectation of the ultrapower were stable, then, by the theorem of Iovino recalled in Variatio 5, it would factor through the pairing of a reflexive space X with its dual, by bounded maps on the two sorts; the closed span in X^* of the images of the countably many probes would then be a separable reflexive space with a separable dual, and the restrictions to this span of the images of the $2^{\aleph_0}$ preparations would be $2^{\aleph_0}$ points of its dual at mutual distances bounded below, which is impossible.

Example 7.2 (Stable but not uniformly stable). Let both systems carry the Hilbert space $\bigoplus_{n\geq 1}\mathbb{C}^n$, whose summands we call *sectors*; the sector $\mathbb{C}^n$ is the Hilbert space of a particle of spin $(n-1)/2$, and only its orthonormal basis $e_1^n,\dots,e_n^n$ is used. Let C_{sec} be the orthogonal projection onto the closed span of the vectors $e_i^n\otimes e_j^n$ with $i\leq j$, which describes the event "both systems lie in the same sector, and Alice's level does not exceed Bob's". In the sector $\mathbb{C}^n$ the basis vectors form a ladder of length n at margin 1, and the observable C_{sec} is therefore not uniformly stable. Nevertheless, it is stable: the effective observable $T_{C_{\mathrm{sec}}}(\rho)$ depends only on the diagonal entries $\rho_{ii}^n=\langle e_i^n,\rho e_i^n\rangle$ of ρ, and the effective map factors through the space X of sequences $(w_n)_{n\geq 1}$, with $w_n\in\mathbb{C}^n$, normed by $(\sum_n\|w_n\|_1^2)^{1/2}$, where $\|w_n\|_1$ is the sum of the absolute values of the entries of w_n. The first factor sends a trace-class operator τ to the sequence $((\tau_{ii}^n)_{i\leq n})_{n\geq 1}$ of the diagonal entries of τ, with norm at most $\|\tau\|_1$, and the second is the linear map, of norm at most 1, that sends $(w_n)_{n\geq 1}$ to the operator $\bigoplus_n\sum_{j\leq n}(\sum_{i\leq j}w_{n,i})|e_j^n\rangle\langle e_j^n|$. The space X, an ℓ^2-sum of finite-dimensional spaces, is reflexive, so condition (A3) of Theorem 5.1 holds.

The model-theoretic reading. A tree of height n is a binary tree of Shelah for the formula φ in $\mathcal{M}_C$, and the largest height of such trees is Shelah's binary tree rank, which, as Chase and Freitag observed [**CF19**], is the Littlestone dimension of the class of responses, the optimal mistake bound of (B5). Condition (B3) says that φ is stable in the theory of $\mathcal{M}_C$, as in the model-theoretic reading of Variatio 2, and the count in (B4) is a count of φ-types over finite sets of probes, a finite and adaptive form of Rowbottom's original definition of stability, by which a theory is stable in an infinite cardinal λ if there are at most λ types over every set of λ parameters [**Row64**].

For the physicist or operator theorist. Uniform stability is stability of the expectation in every ultrapower of the structure $\mathcal{M}_C$, the analogue for observables of Raynaud's superstability of Banach spaces, the stability of all their ultrapowers [**Ray83**]; by Example 7.2, it is strictly stronger than stability.

For the logician. In the response-query model behind Theorem 7.1, Alice supplies copies of an unknown preparation ρ chosen from a known list, and Bob chooses a probe σ at each step. The answer records whether $\varphi_C(\rho,\sigma)\geq t+\delta/2$ or $\varphi_C(\rho,\sigma)\leq t-\delta/2$, with t fixed throughout the protocol. Experimentally, Bob estimates this expectation by measuring C on repeated preparations of $\rho\otimes\sigma$; the theorem counts the queries and assumes their correct threshold labels, rather than bounding the repetitions needed for statistical confidence. Each probe may depend on the previous labels, and a tree of height n distinguishes its 2^n leaf preparations with n such queries. For a uniformly stable observable with bound d on tree height at the prescribed margin, no tree of height greater than d is realized, and Littlestone's learner makes at most d mistakes when predicting the labels. In Example 7.2, the summand $\mathbb{C}^n$ carries the spin-$(n-1)/2$ representation; ever larger multiplets carry ever longer ladders, so qualitative stability does not give this uniform bound.

The ultrapower expectation is a joint limit, not an iterated one, and uniform stability is Grothendieck's double limit condition applied to families of experiments: it holds exactly when trees have bounded height, and then adaptive discrimination grows polynomially and online learning makes boundedly many mistakes.

VARIATIO 8

a 2 Clav.

Bounds independent of the dimension

The swap test and the Hong–Ou–Mandel coincidence (see Variatio 4) are stable in every dimension. In this variation we show that they are uniformly stable, with bounds that do not depend on the dimension. Bob chooses probe states adaptively and queries the coincidence or acceptance probability on fresh copies of Alice's preparation, using the previous response labels to choose the next probe. The bound in Theorem 8.1, combined with the counting argument of Variatio 7, gives a polynomial bound on the number of preparations distinguished by a given number of expectation queries; its degree depends only on the margin. The same bound therefore holds in a polarization model and in a model with infinitely many spatial modes.

Decision trees

An adaptive protocol of Bob, in which each probe is chosen in view of the previous outcomes, is a binary decision tree. Fix a margin $\delta > 0$ and a bounded real function φ on pairs consisting of a preparation of Alice and a probe of Bob, such as the expectation of a bipartite observable. A *tree of height n at margin δ*, in this variation, differs from a tree of Variatio 7 in having a threshold at each node: it assigns a probe σ_s and a threshold t_s to every node s of the complete binary tree of height n, and a preparation ρ_b of Alice to every branch b, in such a way that, whenever the branch b passes through the node s,

$$\varphi(\rho_b, \sigma_s) \geq t_s + \tfrac{\delta}{2} \ \text{ if } b \text{ turns right at } s, \qquad \varphi(\rho_b, \sigma_s) \leq t_s - \tfrac{\delta}{2} \ \text{ if } b \text{ turns left.}$$

The 2^n preparations at the branches are then told apart by n probes, each answered with margin δ. Every tree of Variatio 7 becomes a tree in this sense, with all thresholds equal, once left and right are exchanged at every node if necessary; a bound on the height of the trees of this variation therefore bounds the height of the trees of Variatio 7. Since both answers occur at every node, every threshold satisfies $|t_s| \leq \sup|\varphi|$; we use this bound in the proof below.

Theorem 8.1 (Novikoff [Nov62]).
Suppose that $\varphi(\rho, \sigma) = \langle u(\rho), v(\sigma)\rangle$ for all preparations ρ and probes σ, where u and v are maps into a real Hilbert space $\mathcal{K}$, and that R_u and R_v are positive reals with $\|u(\rho)\| \leq R_u$ and $\|v(\sigma)\| \leq R_v$ for all ρ and σ. Then every tree at margin δ has height at most $16(R_u R_v/\delta)^2$. For the swap test and for the Hong–Ou–Mandel coincidence, with the margin measured in probability, every tree has height at most $4/\delta^2$.

Proof. The idea is a comparison of two rates of growth: a vector corrected once at each step of a walk down the tree grows in length at most like the square root of the number of steps, while its inner product with a fixed vector determined by the branch followed grows at least linearly; the Cauchy–Schwarz inequality, which compares the two, bounds the height of the tree. Write $x_s = (v(\sigma_s), -t_s/R_u)$ in the Hilbert space $\mathcal{K} \oplus \mathbb{R}$; each of these vectors

satisfies $\|x_s\|^2 \le 2R_v^2$. We walk down the tree from the root, keeping a vector w, initially 0. At the node s, we turn left if $\langle w, x_s\rangle \ge 0$ and right otherwise, set $y_s = +1$ for a right turn and $y_s = -1$ for a left one, and replace w by $w + y_s x_s$. Since $y_s\langle w, x_s\rangle \le 0$ at every step, each step increases $\|w\|^2$ by at most $\|x_s\|^2 \le 2R_v^2$, and after n steps the vector w satisfies $\|w\|^2 \le 2nR_v^2$. Let b be the branch that was followed, and let $w^* = (u(\rho_b), R_u)$; then $\|w^*\|^2 \le 2R_u^2$ and $\langle w^*, x_s\rangle = \varphi(\rho_b, \sigma_s) - t_s$. By the definition of a tree, the inequality $y_s\langle w^*, x_s\rangle \ge \delta/2$ holds at every node of b, whence $\langle w, w^*\rangle \ge n\delta/2$. The Cauchy–Schwarz inequality now gives $n\delta/2 \le \sqrt{2n}\, R_v \cdot \sqrt{2}\, R_u$, that is, $n \le 16(R_u R_v/\delta)^2$. For the swap test, the overlap $\operatorname{Tr}\rho\sigma$ is the Hilbert–Schmidt inner product of two operators of norm at most 1, and a margin δ in probability is a margin 2δ in $\operatorname{Tr}\rho\sigma$; the bound becomes $16/(2\delta)^2 = 4/\delta^2$. The same holds for coincidences. □

At margin $\delta = 0.1$ in probability, no protocol of 401 beam-splitter probes tells apart 2^{401} photon states, however many modes the photons have. No such bound holds for arbitrary observables: for the energy comparison of two oscillators, Bob's probe $|l\rangle_B$, his energy eigenstate of level l, answers "yes" exactly on Alice's energy eigenstates $|k\rangle_A$ with $k < l$, and binary search on the levels builds trees of every height at margin 1.

The model-theoretic reading. At a fixed threshold, the largest height of the trees of Variatio 7 is Shelah's binary tree rank of the formula φ in the structure $\mathcal{M}_C$, the Littlestone dimension of the thresholded responses. The trees of this variation allow a threshold at each node, so Theorem 8.1 bounds the rank uniformly in the threshold, and the expectation of an overlap measurement is stable in the theory of $\mathcal{M}_C$, by Theorem 7.1.

For the physicist or operator theorist. The proof of Theorem 8.1 is the mistake bound of the perceptron, a learner that predicts the side of the threshold by the sign of a linear functional on the Hilbert space through which the expectation factors; the bound depends only on the norms of the two factors and on the margin, never on the dimension.

For the logician. The internal space K of a photon is $\mathbb{C}^2$ in a polarization-only model; allowing spatial or temporal modes leads to infinite-dimensional models. For the Hong–Ou–Mandel case of Theorem 8.1, each query is realized by the arrangement of Variatio 4, repeated on fresh copies: Alice's photon has internal state ρ, Bob supplies a photon in a chosen state σ_s, and the ideal lossless balanced beam splitter, preserving the internal modes, gives coincidence probability $\frac{1}{2}(1-\operatorname{Tr}(\rho\sigma_s))$. At margin δ in this probability, a full tree of such expectation queries has height at most $4/\delta^2$, independently of the dimension of K. Repeated runs estimate each probability; they are not counted by the tree height. In each fixed finite dimension some bound follows from compactness of the state space; the bound proved here remains uniform as the dimension grows.

Since a ladder with 2^h rungs contains a tree of height h, by the binary search of Variatio 7, a ladder of the swap test or of the Hong–Ou–Mandel coincidence at margin δ in probability has fewer than $2^{1+4/\delta^2}$ rungs. These overlap measurements therefore have no infinite ladder, and their two orders of idealization agree, by Theorems 2.1 and 3.1: at margin δ in probability, the staircase of the Aria, the pattern of ones above the diagonal and zeros on and below it, can be imitated by the values of an overlap measurement only with fewer than $2^{1+4/\delta^2}$ steps.

VARIATIO 9

a 1 Clav. Canone alla Terza

Mixtures simulate idealized preparations

An oscillator excited to infinite energy cannot be prepared in a laboratory, but a finite mixture of actual preparations, which produces each of them with a prescribed probability, can. In this variation we show that, for stable observables, finite mixtures of actual preparations simulate idealized ones uniformly over all probes, idealized probes included, in the order of idealization with Bob's system idealized first. The approximating mixture can be drawn from any tail of any net of actual preparations converging to the state. Under the Hilbert-factorization hypotheses below, the number of terms is bounded by a quadratic function of the reciprocal error, and for the swap test and the Hong–Ou–Mandel coincidence an equal-weight mixture of N terms, repetitions allowed, suffices for every integer $N \geq 1/\varepsilon^2$, whatever the dimension. The logician averages formulas and the physicist averages preparations; since the expectation of an observable is affine in the state, the two averages coincide.

The logician's reading. Let $\varphi(x, y)$ be a formula, and let M be a model of a theory in which φ is stable. In real-valued logic, the definition of a φ-type p over M, the continuous function on the space $S_y(M)$ of types in y over M that extends the map $b \mapsto \varphi(p, b)$ (see Variatio 3), is a uniform limit of continuous combinations of instances $\varphi(a, y)$ with a in M [**BYU10**]. Ben Yaacov observed that, by Mazur's theorem, according to which a convex subset of a normed space has the same weak and norm closures, averages suffice: the definition is a uniform limit of averages $\frac{1}{n}\sum_{i=1}^{n} \varphi(a_i, y)$, with $n \geq 1$ and $a_1, \ldots, a_n$ in M [**BY14**].

The physicist's reading. Let H_1 and H_2 be the Hilbert spaces of Alice's and of Bob's systems, and let C be a bipartite observable on $H_1 \otimes H_2$, with expectation $\varphi_C(\rho, \sigma) = \operatorname{Tr}((\rho \otimes \sigma)C)$ for ρ in the set $\mathfrak{D}(H_1)$ of density operators on H_1 and σ in $\mathfrak{D}(H_2)$. Recall that the effective observable of a preparation ρ of Alice is the unique bounded operator $T_C(\rho)$ on H_2 with $\operatorname{Tr}(\sigma\, T_C(\rho)) = \varphi_C(\rho, \sigma)$ for every $\sigma \in \mathfrak{D}(H_2)$ (see Variatio 1). Let ω_1 be a state of Alice's algebra $B(H_1)$, and let $\otimes_{\mathsf{B}}$ denote the product of states that corresponds to the order of idealization with Bob's system idealized first.

Theorem 9.1.

(a) *The observable C is stable if and only if, for every state ω_1 of Alice's algebra $B(H_1)$ and every $\varepsilon > 0$, some density operator ρ satisfies $|(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) - \omega_2(T_C(\rho))| \leq \varepsilon$ for all states ω_2 of Bob, idealized or not. The density operator ρ can be taken to be a finite mixture of terms from any tail of any net of density operators converging to ω_1.*

(b) *Let $\mathcal{K}$ be a real Hilbert space, let $u\colon \mathfrak{D}(H_1) \to \mathcal{K}$ and $v\colon \mathfrak{D}(H_2) \to \mathcal{K}$ be maps, and suppose that $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma)\rangle$ for all $\rho \in \mathfrak{D}(H_1)$ and $\sigma \in \mathfrak{D}(H_2)$, that some constant R_u satisfies $\|u(\rho)\| \leq R_u$ for all $\rho \in \mathfrak{D}(H_1)$, and that v extends to a bounded linear map, of norm $\|v\|$, from the self-adjoint trace-class operators on H_2, with the trace norm, to $\mathcal{K}$. Then, for every $\varepsilon > 0$ and every positive integer $N \geq (2R_u\|v\|/\varepsilon)^2$, the density operator in (a) can be taken to be an equal-weight mixture*

of N terms, not necessarily distinct, of any tail of any net of density operators converging to ω_1. For the swap test and the Hong–Ou–Mandel coincidence on $K \otimes K$, where K is a Hilbert space, the same holds for every integer $N \geq 1/\varepsilon^2$, whatever the dimension of K.

PROOF. (a) Suppose that C is stable, let ω_1 be a state of $B(H_1)$, let $(\rho_i)_{i \in I}$ be a net of density operators converging to ω_1, let $i_0 \in I$, and let $\varepsilon > 0$. By the proof of Theorem 5.1 in section P.1, the effective map T_C, extended linearly to the trace-class operators on H_1, is weakly compact, and the bitranspose of T_C sends ω_1 to the operator Z_{ω_1} of Variatio 1. By Gantmacher's theorem [**Gan40**], this bitranspose takes values in $B(H_2)$ and is continuous from the weak* topology of $B(H_1)^*$ to the weak topology of $B(H_2)$. The state ω_1 lies in the weak* closure of the convex hull of the density operators ρ_i with $i \geq i_0$, and the operator Z_{ω_1} therefore lies in the weak closure of the convex hull of their effective observables $T_C(\rho_i)$. By Mazur's theorem [**Maz33**], a convex set has the same weak and norm closures, so some mixture $\rho = \sum_{k=1}^{m} t_k \rho_{i_k}$, with $i_1, \ldots, i_m \geq i_0$, satisfies $\|Z_{\omega_1} - T_C(\rho)\| \leq \varepsilon$. Since C is stable, we have $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = (\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = \omega_2(Z_{\omega_1})$ for every state ω_2 of $B(H_2)$, and this value lies within ε of $\omega_2(T_C(\rho))$.

Conversely, suppose that, for a state ω_1 of $B(H_1)$ and every positive integer n, a density operator ρ_n satisfies $|(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) - \omega_2(T_C(\rho_n))| \leq 1/n$ for every state ω_2 of $B(H_2)$. The operators $T_C(\rho_n)$ are self-adjoint, and the norm of a self-adjoint operator A is the supremum of $|\omega(A)|$ over the states ω; hence $\|T_C(\rho_n) - T_C(\rho_m)\| \leq 1/n + 1/m$, and the sequence $(T_C(\rho_n))_{n \geq 1}$ converges in norm to a self-adjoint operator Z with $\omega_2(Z) = (\omega_1 \otimes_{\mathsf{B}} \omega_2)(C)$ for every state ω_2. At an actual state of Bob, given by a density operator σ, the two products agree (see Variatio 1), so $\operatorname{Tr}(\sigma Z) = \operatorname{Tr}(\sigma Z_{\omega_1})$ for every σ, whence $Z = Z_{\omega_1}$ and $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \omega_2(Z_{\omega_1}) = (\omega_1 \otimes_{\mathsf{A}} \omega_2)(C)$ for every state ω_2. If the approximation holds for every state ω_1, the two products therefore agree at every pair of states, and C is stable by Theorem 3.1.

(b) The argument is Maurey's empirical method [**Pis81**]: Mazur's theorem provides a finite mixture, and random sampling replaces it by an equal-weight mixture of N of its terms. For $\eta \in \mathcal{K}$, the functional $\tau \mapsto \langle \eta, v(\tau) \rangle$ on the self-adjoint trace-class operators on H_2 is bounded and linear, hence of the form $\tau \mapsto \operatorname{Tr}(\tau B_\eta)$ for a unique self-adjoint operator B_η on H_2. The map $\eta \mapsto B_\eta$ is linear, satisfies $\|B_\eta\| \leq \|v\| \, \|\eta\|$, and sends $u(\rho)$ to $T_C(\rho)$ for every $\rho \in \mathfrak{D}(H_1)$, because $\operatorname{Tr}(\sigma B_{u(\rho)}) = \langle u(\rho), v(\sigma) \rangle = \operatorname{Tr}(\sigma \, T_C(\rho))$ for every $\sigma \in \mathfrak{D}(H_2)$.

Let $(\rho_i)_{i \in I}$ be a net in $\mathfrak{D}(H_1)$ converging to ω_1, let $i_0 \in I$ be given, and let ξ^* be a weak cluster point of the net $(u(\rho_i))_{i \in I}$, which exists because closed balls of a Hilbert space are weakly compact. For every state ω_2 of Bob, the inner limit in the definition of $\otimes_{\mathsf{B}}$ is $\omega_2(T_C(\rho_i))$, whence $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \lim_i \omega_2(B_{u(\rho_i)})$. The functional $\eta \mapsto \omega_2(B_\eta)$ on $\mathcal{K}$ is bounded and linear, hence weakly continuous, and passing to a subnet along which $u(\rho_i)$ converges weakly to ξ^* gives $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \omega_2(B_{\xi^*})$. As a cluster point of the net, the vector ξ^* lies in the weak closure of the convex hull of the vectors $u(\rho_i)$ with $i \geq i_0$, hence, by Mazur's theorem, in the norm closure of the same convex hull: there is a convex combination $\xi = \sum_{k=1}^{m} t_k u(\rho_{i_k})$, with indices $i_1, \ldots, i_m \geq i_0$, such that $\|v\| \, \|\xi^* - \xi\| \leq \varepsilon/2$.

Let $\kappa_1, \ldots, \kappa_N$ be independent random elements of $\{1, \ldots, m\}$, each equal to k with probability t_k, and let $\bar{\xi}$ be the average of the vectors $u(\rho_{i_{\kappa_1}}), \ldots, u(\rho_{i_{\kappa_N}})$. These vectors are independent, with mean ξ and norm at most R_u; since the variance of a random vector is at most its mean square norm, we have $\mathbb{E}\|\bar{\xi} - \xi\|^2 = \frac{1}{N}\mathbb{E}\|u(\rho_{i_{\kappa_1}}) - \xi\|^2 \leq R_u^2/N$. Hence some outcome of the draw satisfies $\|\bar{\xi} - \xi\| \leq R_u/\sqrt{N}$, and the choice of N gives $\|v\| R_u/\sqrt{N} \leq \varepsilon/2$. Fix such an outcome, and let $\bar{\rho}$ be the equal-weight mixture of the N density operators $\rho_{i_{\kappa_1}}, \ldots, \rho_{i_{\kappa_N}}$,

which are terms of the net with indices at least i_0. Because T_C is affine on $\mathfrak{D}(H_1)$ and the map $\eta \mapsto B_\eta$ is linear, the operator $T_C(\bar{\rho})$ equals $B_{\bar{\xi}}$. The map u is evaluated only at terms of the net, never at a mixture, and it therefore need not be affine. For every state ω_2 of Bob, it follows that $|(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) - \omega_2(T_C(\bar{\rho}))| = |\omega_2(B_{\xi^* - \bar{\xi}})| \leq \|v\| \, (\|\xi^* - \xi\| + \|\xi - \bar{\xi}\|) \leq \varepsilon$. For the swap F of Lemma 4.1, the unitary on $K \otimes K$ that exchanges the two factors, take $H_1 = H_2 = K$, and let u and v both be the inclusion of $\mathfrak{D}(K)$ into the real Hilbert space of self-adjoint Hilbert–Schmidt operators on K, with the inner product $\langle \tau, \tau' \rangle = \mathrm{Tr}(\tau\tau')$. By the swap trick, these maps satisfy $\varphi_F(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$, and since the Hilbert–Schmidt norm is at most the trace norm, the vectors $u(\rho)$ have norm at most $R_u = 1$ and the extension of v has norm at most 1; every integer $N \geq 4/\varepsilon^2$ therefore suffices for F. The swap test and the Hong–Ou–Mandel coincidence are, by the same lemma, the projections $\frac{1}{2}(I \pm F)$, for which $T_{\frac{1}{2}(I \pm F)}(\rho) = \frac{1}{2}(I \pm T_F(\rho))$ and $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(\frac{1}{2}(I \pm F)) = \frac{1}{2}(1 \pm (\omega_1 \otimes_{\mathsf{B}} \omega_2)(F))$. Their errors are therefore half those of F, and the bound for F at accuracy 2ε shows that every integer $N \geq 4/(2\varepsilon)^2 = 1/\varepsilon^2$ suffices for them, whatever the dimension of K. □

The logician's reading. The order of idealization in part (a) of Theorem 9.1, with Bob's system idealized first, is the order in which a definition is evaluated: Bob's type is substituted into the definition of Alice's.

The physicist's reading. The order cannot be dropped: let $(\psi_k)_{k \geq 0}$ and $(\chi_l)_{l \geq 1}$ be orthonormal bases of H_1 and of H_2, let S_k, for $k \geq 1$, be the orthogonal projection onto the span of $\chi_1, \ldots, \chi_k$, and let $C' = |\psi_0\rangle\langle\psi_0| \otimes I + \sum_{k \geq 1} |\psi_k\rangle\langle\psi_k| \otimes S_k$. Then, for $k \geq 1$, the value $\varphi_{C'}(\psi_k, \chi_l)$ is 1 exactly when $l \leq k$, whence the two iterated limits of these values differ, and the observable C' is unstable. Nevertheless, let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$, and let $\omega_1 = \lim_{k,\mathcal{U}} \psi_k$. With Alice's system idealized first, the actual state ψ_0 reproduces ω_1 exactly: its effective observable $T_{C'}(\psi_0)$ is I, and the response of ω_1 also takes the value 1 at every actual probe σ, because $\mathrm{Tr}(\sigma S_k)$ tends to $\mathrm{Tr}\, \sigma = 1$ as k grows. With Bob's system idealized first, at the idealized probe $\omega_2 = \lim_{l,\mathcal{U}} \chi_l$, the idealized state ω_1 answers 0 and the actual state ψ_0 answers 1. More generally, with Alice's system idealized first, every state ω_1 of $B(H_1)$ is reproduced exactly by some actual state: the effective observables $T_{C'}(\rho)$ of density operators ρ are exactly the operators diagonal in the basis $(\chi_l)_{l \geq 1}$ whose diagonals are nonincreasing sequences in $[0, 1]$ with first term 1; this set is weak* closed, and the operator Z_{ω_1}, which lies in its weak* closure, is therefore $T_{C'}(\rho)$ for some density operator ρ. Since the two orders agree on actual probes, the example also shows that approximation uniform over actual probes alone does not characterize stability. For the energy comparison P of Variatio 1, the failure is total: no mixture comes within any error less than 1, even on actual probes. Indeed, for the state $\omega_1 = \lim_{k,\mathcal{U}} |k\rangle_A$ and the actual probe $|l\rangle_B$, the response of ω_1 is 0, since Alice's level eventually exceeds l, whereas, for a density operator ρ of Alice, the value ${}_B\langle l|T_P(\rho)|l\rangle_B = \sum_{k<l} {}_A\langle k|\rho|k\rangle_A$ tends to 1 as l grows.

Both readings. In general, the definitions of types in real-valued logic are uniform limits of polynomials, or of lattice combinations, of instances. For an observable, an average of instances is again an instance: for density operators $\rho_1, \ldots, \rho_m$ of Alice and weights $p_1, \ldots, p_m \geq 0$ with sum 1, the expectation, being affine, satisfies $\sum_{k=1}^m p_k \varphi_C(\rho_k, \cdot) = \varphi_C(\sum_{k=1}^m p_k \rho_k, \cdot)$. The approximating formula is therefore the instance of a physical preparation, a finite mixture, and polynomials of degree one suffice.

When Grothendieck's double limit condition holds, it places the observable that represents an idealized response of Alice in the weak closure of the effective observables of her actual

preparations; Mazur's theorem then turns the weak limit into a norm limit, and every idealized preparation is thereby approximated, uniformly over all probes and with Bob's system idealized first, by a finite mixture that a laboratory can prepare.

VARIATIO 10

a 1 Clav. Fughetta

Countable randomization does not detect instability

Although no laboratory prepares an oscillator of infinite energy, the level of an oscillator can be drawn at random. Alice draws the level of her oscillator from a probability distribution, Bob draws the level of his from another, and the energy comparison of the two oscillators, which asks whether Bob's level is the higher, is averaged over both draws. The two averages can be taken in either order, Alice's first or Bob's first. In this variation we show that the two iterated averages agree for every observable when the distributions are countably additive, and that they agree for all finitely additive means exactly when the observable is stable.

We compare the two iterated integrals for three classes of measures. It is convenient to integrate over all density operators, not only over energy eigenstates: a *mean* on the set $\mathfrak{D}(H)$ of density operators on a Hilbert space H is a finitely additive probability measure defined on all subsets of $\mathfrak{D}(H)$, and the *barycenter* of a mean μ is the state $A \mapsto \int \operatorname{Tr}(\rho A)\, d\mu(\rho)$ of $B(H)$. Let C be a bipartite observable on $H_1 \otimes H_2$, where H_1 and H_2 are the spaces of Alice's and of Bob's systems, and let μ and ν be means on $\mathfrak{D}(H_1)$ and on $\mathfrak{D}(H_2)$, with barycenters ω_μ and ω_ν. Since the expectation $\varphi_C(\rho,\sigma) = \operatorname{Tr}((\rho\otimes\sigma)C)$ is affine in each state, an iterated integral of φ_C depends only on the barycenters, and it is the value at C of one of the two products of the barycenters, $\otimes_{\mathsf{A}}$ with Alice's system idealized first and $\otimes_{\mathsf{B}}$ with Bob's (see Variatio 1):

$$\int\Big(\int \varphi_C\, d\mu\Big) d\nu = (\omega_\mu \otimes_{\mathsf{A}} \omega_\nu)(C), \qquad \int\Big(\int \varphi_C\, d\nu\Big) d\mu = (\omega_\mu \otimes_{\mathsf{B}} \omega_\nu)(C).$$

Finite mixtures. If μ is finitely supported, its barycenter is a mixture, a density operator, and finite sums commute. The two integrals agree for every observable.

Countably additive measures. Let μ be a countably additive Borel probability measure on $\mathfrak{D}(H_1)$, with the trace norm. Since H_1 is separable, so is the space of trace-class operators on H_1, and the identity map $\rho \mapsto \rho$ is strongly measurable and has constant trace norm one. Its Bochner integral $\bar\rho = \int \rho\, d\mu(\rho)$ therefore exists, is positive and has trace one, and represents the normal barycenter of μ. Both iterated integrals are defined even when ν is only a mean: averaging the functions $\rho \mapsto \varphi_C(\rho,\sigma)$ against ν preserves their common Lipschitz constant $\|C\|$, so the outer integrand for μ is Borel measurable; the outer integral against ν requires only a bounded integrand. The two products agree whenever one factor is actual, so the two iterated integrals agree for every observable and every mean ν of Bob. No countably additive randomization detects instability, for any observable.

Finitely additive means. For finitely additive means, the answer is stability (Theorem 10.1 below).

Theorem 10.1 (Pták [Ptá63], Simons [Sim68]). *An observable C is stable if and only if, for all means μ on $\mathfrak{D}(H_1)$ and ν on $\mathfrak{D}(H_2)$,*

$$\int\Big(\int \varphi_C(\rho,\sigma)\,d\mu(\rho)\Big)d\nu(\sigma) = \int\Big(\int \varphi_C(\rho,\sigma)\,d\nu(\sigma)\Big)d\mu(\rho).$$

Proof. The two sides are the two products of the barycenters. Every state is a barycenter: if ω is the limit of a net (ρ_i) of density operators, and $\mathcal{W}$ is an ultrafilter on the index set containing all tails, then the mean μ given by $\mu(S) = \lim_{i,\mathcal{W}} \mathbf{1}_S(\rho_i)$, for $S \subseteq \mathfrak{D}(H)$, has barycenter ω. Hence the identity holds for all means exactly when the two orders of idealization agree for all states, that is, by Theorem 3.1, exactly when C is stable. □

For the energy comparison, let μ and ν be "uniform distributions on all levels": invariant means on $\mathbb{N}$, such as Banach limits of the sequence of uniform distributions on $\{1,\dots,n\}$, for $n \geq 1$, placed on the energy eigenstates $|k\rangle_A$ of Alice's oscillator and $|l\rangle_B$ of Bob's, for $k,l \in \mathbb{N}$. For each fixed l, almost every k in the sense of μ exceeds l, so the inner integral over Alice's levels is 0; for each fixed k, almost every l in the sense of ν exceeds k, so the inner integral over Bob's levels is 1. The two iterated integrals are 0 and 1. The same happens for ultrafilters, the $\{0,1\}$-valued means: for a nonprincipal ultrafilter $\mathcal{W}$ on $\mathbb{N}$, the set $\{(k,l) : k < l\}$ has measure 0 under $\lim_{l,\mathcal{W}} \lim_{k,\mathcal{W}}$ and measure 1 under $\lim_{k,\mathcal{W}} \lim_{l,\mathcal{W}}$. No countably additive version exists, because a countably additive distribution on $\mathbb{N}$ cannot assign measure zero to every finite set.

The model-theoretic reading. A mean μ on Alice's state space $\mathfrak{D}(H_1)$ defines a Keisler measure over the structure $\mathcal{M}_C$ of Variatio 1, the functional $\theta(x,b) \mapsto \int \theta(\rho,b)\,d\mu(\rho)$ on formulas with parameters from $\mathcal{M}_C$; a mean that takes only the values 0 and 1 is given by an ultrafilter on $\mathfrak{D}(H_1)$, and its Keisler measure is the type of an element of an ultrapower. The two iterated integrals of Theorem 10.1 are the two products of Keisler measures, the averaged form of the two independent amalgams of Variatio 6, and the theorem says that these products commute for all measures exactly when they commute for all types. A countably additive measure on $\mathfrak{D}(H_1)$ with the trace norm has a density operator as barycenter, an element of $\mathcal{M}_C$ itself; for the formula φ it behaves like a realized type, and the two products agree because they agree whenever one factor is realized.

For the physicist or operator theorist. The theorem is Fubini's theorem for finitely additive probability measures on preparations, which logicians call Keisler measures; ultrafilters give the extreme case, the types, and countably additive measures satisfy the theorem because their barycenters are density operators.

For the logician. A laboratory that draws Alice's preparation at random, by a countably additive distribution μ, and does not record the draw, prepares the density operator $\int \rho\,d\mu(\rho)$, since the expectation of every observable is affine in the state. Hence, if either system is prepared by a countably additive randomization, this system is in an actual state, and the two iterated averages agree for every observable, because the two orders of idealization agree when one factor is actual (see Variatio 1). Any disagreement of the two iterated averages, such as the one found for the energy comparison, therefore needs means on both systems that are not countably additive, for example means on the levels that vanish on every finite set; the barycenter of such a mean is an idealized state, given by no density operator, that no random procedure produces.

Grothendieck's double limit condition is thus a Fubini theorem. For finite mixtures and countably additive measures it holds for every observable; for finitely additive means it holds exactly for the stable ones.

Part II. NIP

VARIATIO 11

a 2 Clav.

The independence property

Photons that pass a double slit are detected on Bob's screen in bright and dark fringes, whose spacing is inversely proportional to the wavenumber of the photons, 2π divided by their wavelength. If photons are prepared with wavenumbers $8, 8^2, 8^3, \dots$, measured in units of the reciprocal of a fixed length, then, as we show in Variatio 16, a single position of Bob's detector can be bright for any chosen finite set of them and dark for any other finite set disjoint from it: the fringes realize every pattern. The energy comparison of the two oscillators of the Aria cannot do this, because its answers are monotone. Both measurements are unstable; Part II is devoted to the difference between their two kinds of instability. In this variation we draw the second dividing line, between the instability of the energy comparison, which is NIP, and instability with the independence property, and name the configuration that NIP forbids.

A family of real functions on a set Y has the *independence property* if there are a sequence $(r_n)_{n\in\mathbb{N}}$ in it and fixed reals $a < b$ such that every finite pattern is realized: for all disjoint finite sets $E, F \subseteq \mathbb{N}$, some $y \in Y$ satisfies $r_n(y) \le a$ for $n \in E$ and $r_n(y) \ge b$ for $n \in F$. For a bipartite observable C with expectation φ_C, the *independence property for Alice's responses* is this property of the functions $\sigma \mapsto \varphi_C(\rho, \sigma)$, where ρ ranges over Alice's preparations and σ over Bob's probes, and the absence of the property is *NIP for Alice's responses*; unqualified NIP will refer to this direction. In Theorem 11.1 below, we show that stable observables are NIP; the energy comparison is NIP, since its answers are monotone, but unstable.

The direction matters, because NIP for an observable is a condition on a single structure. For a formula in a complete theory, NIP concerns finite configurations in all models of that theory and is symmetric under exchange of the variables [**She90**]. The counterpart for an observable is uniform NIP, which bounds, at each positive gap $b - a$, the sizes of the finite sets of responses on which probes realize every pattern, and which is symmetric in Alice and Bob. NIP forbids only an infinite independent sequence in the given structure, and need not be symmetric: a spin chain read site by site is NIP for Alice's responses but not for Bob's (see Variatio 14 and Corollary P.52).

The responses in the theorem below are functions on a space of probes. Write $\mathfrak{A}_C$ for the unital C*-algebra generated by the effective observables $T_C(\rho)$ of Alice's preparations ρ (see Variatio 1), the smallest algebra of operators on Bob's Hilbert space that contains them and the identity and is closed under adjoints and norm limits, and call it the *probe algebra*; write $Y_C = \mathcal{S}(\mathfrak{A}_C)$ for its state space, with the weak* topology, and call it the *probe space*. Since H_1 is separable, so is $\mathfrak{A}_C$, and Y_C is compact and metrizable (see Lemma P.22). For each $\sigma \in \mathfrak{D}(H_2)$, its *probe record* is the state $\mathfrak{p}_\sigma \colon \mathfrak{A}_C \to \mathbb{C}$ defined by $\mathfrak{p}_\sigma(a) = \operatorname{Tr}(\sigma a)$. For each $\rho \in \mathfrak{D}(H_1)$, define the *completed response* $\mathfrak{r}_\rho \colon Y_C \to [-\|C\|, \|C\|]$ by $\mathfrak{r}_\rho(\omega_2) = \omega_2(T_C(\rho))$ for $\omega_2 \in Y_C$. It recovers the response on actual probes through $\mathfrak{r}_\rho(\mathfrak{p}_\sigma) = r_\rho(\sigma)$. Fraktur indicates the domain of the response, not whether the preparation or the probe is idealized. The completed responses are continuous on Y_C, and we call the members of their pointwise

closure, the completed responses included, the *idealized responses*; they are the responses of the states of $B(H_1)$ with Bob's system idealized first (see Lemma P.24). Restriction to $\mathfrak{A}_C$ maps the state space of $B(H_2)$ continuously onto Y_C (see the proof of Lemma P.22); composition with this map carries the completed responses and the idealized responses onto the responses and the idealized responses of the Aria, and a function on Y_C is continuous exactly when its composite is.

In the theorem below, a function is of the *first Baire class* if it is a pointwise limit of a sequence of continuous functions, a *Rosenthal compactum* is a set of such functions on a Polish space that is compact in the topology of pointwise convergence, and $\beta\mathbb{N}$ is the Stone–Čech compactification of $\mathbb{N}$, whose points are the ultrafilters on $\mathbb{N}$.

Theorem 11.1. *Let C be a bipartite observable.*

(a) *The observable C is stable if and only if every idealized response is continuous on Y_C.*

(b) *The observable C is NIP if and only if every idealized response is of the first Baire class. The idealized responses then form a Rosenthal compactum.*

(c) *If C is not NIP, then the idealized responses contain a homeomorphic copy of $\beta\mathbb{N}$.*

Parts (b) and (c) of the theorem are Rosenthal's dichotomy, in the form extended by Bourgain, Fremlin and Talagrand [**Ros74**, **BFT78**], applied to the completed responses on Y_C; part (a) is the counterpart on Y_C of Grothendieck's theorem as read in the Aria. The proof is in section P.2.

The model-theoretic reading. In a theory, the independence property of a formula, the existence for every n of n parameters that admit all 2^n patterns, yields by compactness an infinite independent sequence in an elementary extension; the definition above instead asks for an infinite independent sequence in the single structure $\mathcal{M}_C$ of Variatio 1. For a sequence $(\rho_n)_{n\in\mathbb{N}}$ of Alice's preparations and a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, let $\mathfrak{r} = \lim_{n,\mathcal{U}} \mathfrak{r}_{\rho_n}$, with the limit taken pointwise on Y_C. Its restriction $\sigma \mapsto \mathfrak{r}(\mathfrak{p}_\sigma)$ to actual probes is the φ-type of $[\rho_n]_{\mathcal{U}}$ over $\mathcal{M}_C$, as described in Variatio 1. The correspondence concerns the formula φ alone, and it concerns the coheir obtained from the same sequence along the same ultrafilter; other coheirs of the same restriction can take other values at idealized probes. An element of Bob's sort in an elementary extension is represented by a probe state, a point ω_2 of the compact metrizable space Y_C with $\omega_2(T_C(\rho)) = \varphi(\rho, b)$ for every actual preparation ρ of Alice, where b is the element; such a state exists, by elementarity and the compactness of Y_C. Two representing states agree at every completed response $\mathfrak{r}_\rho$ of an actual preparation, hence on every pointwise limit of such responses, so the value of the selected coheir at the element is the value of $\mathfrak{r}$ at any representing state. Complete types are not identified with states, nor completed records with density operators. Distinct idealized responses can have the same restriction, so the topology in Theorem 11.1 concerns functions on Y_C, not merely their actual-probe restrictions. For NIP observables, every idealized response is the pointwise limit on Y_C of a sequence of completed responses of actual preparations; Simon proved the corresponding sequential result for NIP formulas [**Sim15b**]. An independent sequence of preparations instead yields a copy of $\beta\mathbb{N}$ in the space of idealized responses.

For the physicist or operator theorist. By Rosenthal's ℓ^1 theorem [**Ros74**], the observable C is NIP exactly when no sequence of effective observables $T_C(\rho_n)$ of Alice's preparations

is equivalent, in the operator norm, to the unit vector basis of ℓ^1: a bounded sequence of continuous functions on the compact space Y_C has either a pointwise convergent subsequence or a subsequence equivalent to that basis in the supremum norm, and on self-adjoint elements of $\mathfrak{A}_C$ the supremum norm on Y_C is the operator norm.

For the logician. In the double-slit experiment Bob measures the position at which the photon is detected, and the response of a preparation of wavenumber k to a detector at the position x of a distant screen, in units in which the two paths to the detector differ in phase by kx, is the relative intensity $\frac{1}{2}(1+\cos kx)$ of the fringes there. The independence property of the fringes is a statement about single detector positions: for the wavenumbers $8, 8^2, 8^3, \ldots$, the intensities at x are close to 1 or close to 0 according to the base-8 digits of $x/2\pi$, which can be chosen freely (see Variatio 16). For each finite pattern, continuity allows the point probe to be replaced by an actual preparation localized near x, for which the averaged relative intensities are at least $\frac{5}{8}$ on the high indices and at most $\frac{3}{8}$ on the low indices. The gap is therefore independent of the finite pattern. For the comparison of photon numbers, by contrast, which is the energy comparison of the Aria when the two oscillators are modes of light, Bob's state $|l\rangle_B$, with l photons, answers yes exactly for Alice's states $|k\rangle_A$ with $k < l$ photons, so on any n levels only the $n+1$ initial segments occur as patterns. When a spin chain is scanned site by site (see Variatio 14), the roles are reversed: the chains of Alice realize every pattern on Bob's sites, which is why NIP for Alice's responses does not imply NIP for Bob's.

When the double limit condition fails, there are two cases. If the observable is NIP, every idealized response is the limit of a sequence of responses of actual preparations; otherwise the idealized responses contain a copy of $\beta\mathbb{N}$, whose nonprincipal points are not limits of such sequences.

VARIATIO 12

Canone alla Quarta

The rational order

The energy comparison of Part I compares two discrete quantities; its continuous counterpart, the comparison of two positions on a line, is governed by the rational order, in which every finite linear order embeds. The rational order has in Part II the role that Hilbert space, the Fraïssé limit of the finite-dimensional Hilbert spaces, has in Part I (see Variatio 4). In this variation we approach the rational order from two directions: the logician descends from the theory of the rational order to its types, and the physicist ascends from two particles on a line to the idealized limits of their comparison.

The finite linear orders form a Fraïssé class, and their limit (see Variatio 4) is $(\mathbb{Q}, <)$, the countable dense linear order without endpoints. Its theory is NIP and unstable, the prototype of a theory that is ordered but tame. The types of the formula $x < y$ over the real line are the cuts, and each real number r splits into two of them, just below and just above r; in the ultrapower $(\mathbb{R}, <)^{\mathcal{U}}$ along a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the class of the sequence $(r + 1/n)_{n\geq 1}$ realizes the type just above r and the class of $(r - 1/n)_{n\geq 1}$ the type just below. The *split interval* of Alexandroff is obtained from the interval $[0, 1]$ by splitting each point into a left and a right copy, with the topology of the lexicographic order; its least and greatest points are isolated. The space of these types is homeomorphic to the *reduced split interval*, the split interval without these two points; the two types at infinity correspond to the remaining copies of 0 and 1. The reduced split interval is hereditarily separable, every subspace having a countable dense subset, but not metrizable. The physical home of the rational order is the comparison of two continuous quantities, such as the positions of two particles on a line.

Counting types measures the complexity of a formula: over finite sets of parameters the count separates NIP from the independence property, and over infinite sets it also separates stability from NIP. Write $\varphi(x, y)$ for the formula $x < y$. Over a finite set of n parameters it has only $n + 1$ φ-types, since a new element falls into one of the $n + 1$ gaps that the parameters leave. A set of n parameters is *shattered* by a formula $\theta(x, y)$ if each subset of the set is cut out by $\theta(a, y)$ for some a, and the *dual Vapnik–Chervonenkis dimension* of θ is the largest n, if there is one, for which some n parameters are shattered; by Łoś's theorem, unbounded shattering in a structure M yields the independence property in the ultrapower $M^{\mathcal{U}}$ along a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, and a bound on shattering is therefore NIP in its finite form. By the Sauer–Shelah lemma, whose tree form appears in Variatio 7, a formula of dual Vapnik–Chervonenkis dimension d has at most $\sum_{i\leq d} \binom{n}{i}$ types over n parameters, a polynomial of degree d in n, and the formula φ, of dimension one, attains the bound $n + 1$, whereas a formula with the independence property has, for every n, 2^n types over some set of n parameters. Over an infinite set the count changes character. The types are the cuts of the set, and their number is governed, for a set of infinite cardinality κ, by the cardinal $\operatorname{ded}\kappa$, the supremum of the numbers of cuts of linear orders of size κ; Shelah showed that,

for a theory T, with $|T|$ the cardinality of its language plus $\aleph_0$, there are, over a set of size κ, at most $\kappa^{|T|}$ complete types if T is stable and at most $(\operatorname{ded}\kappa)^{|T|}$ if T is NIP [**She90**, **CS16**]. The passage from stability to NIP is the passage from points to cuts.

Independence of types changes accordingly. Recall that a type over a monster model of a complete theory is called *global* (see Variatio 1). In an NIP theory, a global type does not fork over a model M, in the sense recalled in Variatio 6, exactly when it is invariant over M, that is, fixed by every automorphism of the monster model that fixes M pointwise [**HP11**]; the coheirs of Variatio 1 are of this kind. Unlike the stable case of Variatio 6, a type over a model may have several invariant extensions. The rational order exhibits two of them for the type just below a real number r. When the type just below r is realized for Alice and for Bob in an iterated ultrapower of the structure $\mathcal{M}_C$ of the comparison C of the positions of two particles on a line (see Variatio 1), the two orders yield two independent amalgams: if Alice's particle, spread uniformly over $(r-1/n, r)$ for $n \geq 1$, is idealized first, it lies closer to r than Bob's, spread uniformly over $(r-1/m, r)$ for $m \geq 1$, and the comparison "Alice's particle lies to the left of Bob's" takes the value 0; in the other order it takes the value 1. This is the staircase of the Aria, compressed into a single point.

The model-theoretic reading. For the formula $x < y$, the ultrapower $(\mathbb{R}, <)^{\mathcal{U}}$ exhibits the types: over $\mathbb{R}$, the element $[r+1/n]$ has the φ-type $y \mapsto 1_{(r,\infty)}(y)$ and the element $[r-1/n]$ the φ-type $y \mapsto 1_{[r,\infty)}(y)$, two functions of the parameter that differ only at $y = r$. Together with the two types of infinite elements, these are all the φ-types over $\mathbb{R}$, and each is a pointwise limit of a sequence of realized types, the behavior that Simon's theorem guarantees for the types of an NIP formula over a countable model [**Sim15b**]. Since the two types of such a pair are separated only by the parameter r itself, no countable set of parameters separates all of them, and the type space, the reduced split interval, is hereditarily separable but not metrizable.

For the physicist or operator theorist. The types of $x < y$ over the line are the indicator functions of the half-lines (r, ∞) and $[r, \infty)$, with the two infinite cases, and the pointwise limits of such indicators are again indicators of half-lines; the space they form is the reduced split interval.

For the logician. For two particles on a line the comparison is the operator of multiplication by the indicator of $\{(x, y) : x < y\}$ on $L^2(\mathbb{R}^2) = L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$, and in states ρ and σ it answers yes with probability $(\mu_\rho \otimes \mu_\sigma)(\{x < y\})$, where μ_ρ and μ_σ are the distributions of the two positions. If Alice's particle is spread over $(r, r+1/n)$ and n grows, its responses converge to $\sigma \mapsto \mu_\sigma((r, \infty))$, and if it is spread over $(r-1/n, r)$, to $\sigma \mapsto \mu_\sigma([r, \infty))$. On actual probes the two limits agree, since a density operator assigns probability zero to the point r; they disagree on the idealized probe obtained by spreading Bob's particle over $(r-1/m, r+1/m)$ and letting m grow first, where the left limit answers 1 and the right limit answers 0. These two idealizations of a particle at r form a split pair.

The rational order thus fails the double limit condition in the mildest way: at each point of the line, the limits of a response from the left and from the right may disagree, these one-sided limits form the split pairs, and every type remains of the first Baire class.

VARIATIO 13

a 2 Clav.

The dictionary of NIP

Simon showed that, for an NIP formula and a countable model, the global types of the formula (see Variatio 12) that are finitely satisfiable in the countable model, viewed as functions on the space of types of parameters over it, are of the first Baire class, each the limit of a sequence of types realized in it [**Sim15b**]. Bipartite observables present the same situation, with Bob's probes as parameters and Alice's idealized responses as types. In the language of Variatio 1, such a type is the response of an idealized preparation of Alice, the limit along an ultrafilter of a sequence of her actual preparations, evaluated, with Bob's system idealized first, at the elements of Bob's sort in any elementary extension; these types are coheirs, hence invariant extensions in the sense of Variatio 12, which do not fork (see Variatio 6). A Rosenthal compactum may be metrizable or very far from it, and the levels of NIP record how far: an NIP observable is $\mathrm{NIP}^{\mathrm{met}}$, $\mathrm{NIP}^{\mathrm{hs}}$ or $\mathrm{NIP}^{\mathrm{fc}}$ when its idealized responses form a metrizable, a hereditarily separable or a first countable compactum, respectively. As in the classification of Dueñez, Iovino, Matos-Wiederhold, Salvetti and Tall [**DIMW$^+$26**, Section 4.2], we write these levels with superscripts rather than numerical subscripts, because in model theory NIP_k denotes Shelah's k-dependence. In this variation we set out the second dictionary (see Table 3) and describe each of these levels by how many probes, and which, are needed to separate the idealized responses.

Two facts about Rosenthal compacta, the sets of functions of the first Baire class on a Polish space that are compact in the topology of pointwise convergence (see Variatio 11), carry the dictionary. The first is the Bourgain–Fremlin–Talagrand theorem that a Rosenthal compactum is *Fréchet–Urysohn*: every point in the closure of a set is the limit of a sequence from the set [**BFT78**]. For a bipartite observable C, recall the probe space Y_C and the completed responses $\mathfrak{r}_\rho \colon Y_C \to [-\|C\|, \|C\|]$ from Variatio 11. Their values at actual probe records satisfy $\mathfrak{r}_\rho(\mathfrak{p}_\sigma) = r_\rho(\sigma) = \varphi_C(\rho, \sigma)$. If C is NIP, then, for every sequence $(\rho_n)_{n\in\mathbb{N}}$ of Alice's preparations and every nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the idealized response $\lim_{n,\mathcal{U}} \mathfrak{r}_{\rho_n}$ is therefore also $\lim_k \mathfrak{r}_{\rho_{n_k}}$ for some sequence of indices $(n_k)_{k\in\mathbb{N}}$, with both limits taken pointwise on Y_C; the ultrafilter, indispensable in Variatio 1, can be replaced by a sequence. For the interference fringes of photons of wavenumbers $8, 8^2, 8^3, \ldots$ behind a double slit (see Variatio 11), this fails: the closure of their responses contains a copy of $\beta\mathbb{N}$, which has no nontrivial convergent sequences, and the points of the copy that correspond to nonprincipal ultrafilters are ultrafilter limits that no sequence reaches. The second fact is that the topology of the closure is pointwise convergence on the whole probe space Y_C, not only at the records of actual probes (see Variatio 11), so each of the levels $\mathrm{NIP}^{\mathrm{met}}$, $\mathrm{NIP}^{\mathrm{hs}}$ and $\mathrm{NIP}^{\mathrm{fc}}$ is a statement about how many probes, and which, decide convergence.

A single example shows both facts at work: the comparison of the positions of two particles on a line (see Variatio 12). Fix a real number r. As n grows, the completed responses of Alice's particle spread uniformly over $(r, r + 1/n)$ converge at every probe, and on actual probes the

Model theory and topology	Quantum measurement
the independence property	a sequence of Alice's preparations on which Bob's probes realize every finite pattern at fixed thresholds: fringes
NIP	comparisons of physical quantities, and the causal order
the first Baire class (Rosenthal; Bourgain, Fremlin, Talagrand)	every idealized response is the limit of a sequence of actual responses
a Rosenthal compactum	the idealized responses of an NIP observable
metrizable, and for observables hereditarily separable	one countable set of probes separates all idealized responses
first countable, not hereditarily separable	each response has its own countable determining probes, but no common countable separating set exists
a split pair	the two pure-phase responses at one transition point in an offset magnetization family
a closed subspace of the Helly space	the idealized responses of a comparison of positions
$\hat{A}(2^{\mathbb{N}})$, not first countable	a light cone shrinks to a point
the seven minimal families	selected trees of preparations of photons, a particle in a box or a spin chain
symmetry of NIP in the two variables	holds for uniform NIP; fails for NIP: a spin chain read site by site

TABLE 3. The dictionary of NIP, read from Alice's preparations to Bob's probes, with Bob's system idealized first.

limit is $\sigma \mapsto \mu_\sigma((r,\infty))$, so a sequence, without an ultrafilter, reaches the idealized response at the cut just above r; the two one-sided idealized responses at r agree on actual probes and are told apart only by idealized probes assigning positive mass to r, with Bob's system idealized first. The profiles of the responses on point probes are monotone in the position of Bob's particle, each with at most countably many jumps, and they are functions of the first Baire class, as Simon's theorem requires.

Metrizable compacta (NIP$^{\mathrm{met}}$). A compact set of functions is metrizable exactly when countably many points separate it (see Proposition P.18). For types, this says that one fixed countable set of instances of the formula decides every limit type. For an observable, this says that countably many probes tell the idealized responses apart. For a mode of light, an oscillator whose energy eigenvectors $e_0, e_1, e_2, \dots$ are the states with no photon, one photon, two photons, and so on, the energy comparison of the Aria is the comparison of photon numbers, which is NIP$^{\mathrm{met}}$, because the idealized responses can jump only at the countably many levels of the spectrum.

Hereditarily separable compacta (NIP$^{\mathrm{hs}}$). A separable Rosenthal compactum is hereditarily separable exactly when it has no discrete subspace of the size of the continuum, by Pol's theorem, which also follows from Todorčević's trichotomy [**Tod99**] (see the proof of Proposition P.19). For types, this says that there is no continuum of limit types each of which finitely many instances tell apart from all the others. The comparison of positions fails this condition. At each point of the line it has a split pair, the two idealized responses of a particle that approaches the point from the right and from the left (see Variatio 12), and the *midpoint* of the pair, the response of the equal mixture of the two idealized preparations, is the average of the two responses. These midpoints, one for each point of the line, form an uncountable discrete set: at the idealized probe concentrated at a point r, the midpoint of the pair at r takes the value $\frac{1}{2}$, and every other midpoint takes the value 0 or 1. For observables, the condition and metrizability coincide (see Variatio 19), which is why NIP$^{\mathrm{hs}}$ and NIP$^{\mathrm{met}}$ merge.

First countable compacta (NIP$^{\mathrm{fc}}$). A point of a compact space of functions has a countable base of neighborhoods exactly when there is a countable set of coordinates on which no other point of the compactum agrees with it (see the proof of Proposition P.19). For types, this says that each limit type is decided by countably many instances chosen for it alone. The comparison of positions is of this kind, NIP$^{\mathrm{fc}}$: a response that jumps at a point is decided by the probes at the point and at rational points around it.

Beyond first countability. The causal order of a $(1+1)$-dimensional spacetime and the comparison of phases on a circle, both defined in Variatio 14, have limit types that no countable set of instances decides. For the causal order, take events of Alice closer and closer to Bob's time slice, whose light cones meet the slice in intervals shrinking to a point c: their idealized limit is distinguished from nearby idealized responses only by probes that assign positive mass to c, and a countable set of probes assigns positive mass to only countably many points.

Read in this way, the classes measure how far an idealized response may be from continuous while remaining a limit of actual ones. Continuity, which Grothendieck's theorem equates with the double limit condition, is the strongest of them: a continuous response is determined at each probe by the probes near it.

VARIATIO 14

a 2 Clav.

Todorčević's trichotomy

When countably many probes no longer separate the idealized responses, the failure takes one of three specific forms [**Tod99**]. In this variation we state Todorčević's trichotomy for separable Rosenthal compacta (see Variatio 11), such as the space of idealized responses of an NIP observable, and list, in Table 4, a physical representative of each class. The table anticipates two later results: in Theorem 16.1 of Variatio 16 we place the representatives in their classes, and in Theorem 19.1 of Variatio 19 we show that the hereditarily separable, non-metrizable class is empty for observables.

Recall that the idealized responses are metrizable exactly when countably many probes separate them (see Proposition P.18). A separable Rosenthal compactum that is not metrizable contains a homeomorphic copy of one of a few canonical compacta: one of the two split Cantor sets $\hat{S}_+(2^{\mathbb{N}})$ and $\hat{S}_-(2^{\mathbb{N}})$ if it is hereditarily separable; the duplicates $\hat{D}(2^{\mathbb{N}})$ or $\hat{D}(S(2^{\mathbb{N}}))$ if it is first countable but not hereditarily separable; and the compactum $\hat{A}(2^{\mathbb{N}})$ if it is not first countable. These compacta keep the names under which Argyros, Dodos and Kanellopoulos introduced them, in which $2^{\mathbb{N}}$ stands for the Cantor set $\{0,1\}^{\mathbb{N}}$ and $2^{<\mathbb{N}}$ for the set $\{0,1\}^{<\mathbb{N}}$ of finite sequences of zeros and ones. Each of them has a countable dense set of isolated points, which the hat records, and is built over a classical space: the split Cantor sets over the split Cantor set $S(2^{\mathbb{N}})$, obtained from the Cantor set $\{0,1\}^{\mathbb{N}}$ by splitting each point into a left and a right copy; the duplicates over the Alexandroff duplicates of the Cantor set and of $S(2^{\mathbb{N}})$, respectively, in which each point receives an isolated twin; and the compactum $\hat{A}(2^{\mathbb{N}})$ over the one-point compactification of a discrete set of cardinality $2^{\aleph_0}$. The decision tree in Figure 4 displays the trichotomy, as in [**DIMW$^+$26**, Section 4.2]; its dashed leaf is absent for full observable compacta (see Theorem 19.1).

The observables of Part II

For the logician, we define once the observables that represent the classes, and the other observables that we classify in Part II. A particle on a σ-finite measure space X with $L^2(X)$ separable is described by $L^2(X)$, and its position is the family of the operators M_{1_E} of multiplication by the indicator functions 1_E of measurable sets $E \subseteq X$: a state ρ places the particle in E with probability $\mu_\rho(E) = \mathrm{Tr}(\rho M_{1_E})$. For particles of Alice and of Bob on two such measure spaces X and Y and a measurable set $R \subseteq X \times Y$, the *relation observable* C_R is multiplication by 1_R on $L^2(X) \otimes L^2(Y) = L^2(X \times Y)$; it represents the event "Alice's value and Bob's value are related by R", and its expectation is $\varphi_{C_R}(\rho, \sigma) = (\mu_\rho \otimes \mu_\sigma)(R)$ (see Lemma P.33). For countable sets A and B, with standard bases $(|a\rangle)_{a\in A}$ of $\ell^2(A)$ and $(|b\rangle)_{b\in B}$ of $\ell^2(B)$, and a function $\varphi\colon A \times B \to [0,1]$, the *diagonal observable* $C_\varphi = \sum_{a\in A,\, b\in B} \varphi(a,b)\, |a\rangle\langle a| \otimes |b\rangle\langle b|$ on $\ell^2(A) \otimes \ell^2(B)$ reads $\varphi(a,b)$ on the pair of labels (a,b).

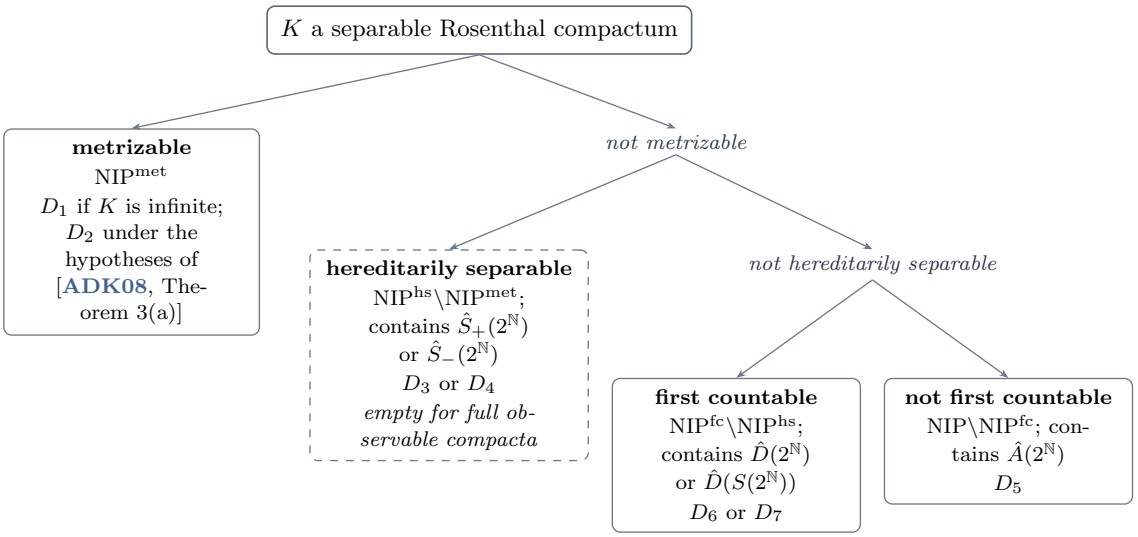


FIGURE 4. Todorčević's trichotomy as a decision tree, with the canonical compacta and the minimal families of the heptachotomy at its leaves.

Class	Witness	Response interpretation	Realization
stable	continuous	idealizations commute	swap test
$\mathrm{NIP}^{\mathrm{met}}\backslash$stable	metrizable	one countable set of probes separates all responses	photon numbers; spin chains read site by site
$\mathrm{NIP}^{\mathrm{hs}}\backslash\mathrm{NIP}^{\mathrm{met}}$	$\hat{S}_+(2^{\mathbb{N}})$ or $\hat{S}_-(2^{\mathbb{N}})$	mixtures add distinguishable midpoints	absent for observables (see Theorem 19.1)
$\mathrm{NIP}^{\mathrm{fc}}\backslash\mathrm{NIP}^{\mathrm{hs}}$	duplicate	each response has its own countable determining probes	positions; quadratures; Ising family with offsets
$\mathrm{NIP}\backslash\mathrm{NIP}^{\mathrm{fc}}$	$\hat{A}(2^{\mathbb{N}})$	some response has no countable determining set of probes	the causal order; phases on a circle
not NIP	$\beta\mathbb{N}$	every finite pattern at a fixed positive gap	interference fringes

TABLE 4. The topology of response compacta and its interpretation through probes.

- *Photon numbers.* A mode of light is an oscillator, as in the Aria, and its energy eigenket $|n\rangle_A$, for $n \in \mathbb{N}$, is the state with n photons; the comparison of photon numbers is the energy comparison P of Variatio 1, the relation observable of $\{(k,l) : k < l\}$ on $\mathbb{N} \times \mathbb{N}$.
- *Positions.* The comparison of positions is the relation observable of $\{(x,y) : x < y\}$ for particles on a line, or in a box $[0,1]$; a particle *spread over* an interval I is in the state $|I|^{-1}|1_I\rangle\langle 1_I|$, where $|I|$ is the length of I. The *quadratures* of a mode are $Q_\theta = Q\cos\theta + P\sin\theta$ on $L^2(\mathbb{R})$ for $\theta \in [0, 2\pi)$, where Q is position, multiplication by q, and P is

momentum, the closure of $-i\,d/dq$ on the rapidly decreasing smooth functions, and where Q_θ stands for the closure of $Q\cos\theta + P\sin\theta$ on the same domain. Each quadrature is self-adjoint and unitarily equivalent to Q, so comparing quadratures is unitarily equivalent to comparing positions.

- *Light cones.* In units in which light has speed 1, an event (t, x) of a $(1+1)$-dimensional spacetime with $t \le 0$ can influence a detection at the point y of Bob's time slice $t = 0$ exactly when $|y - x| \le -t$, that is, when $a \le y \le b$ for the points $a = x + t$ and $b = x - t$ where the two light rays through the event cross the slice. The causal-order observable is the relation observable of $\{(a, b, y) : a \le y \le b\}$ on $[0,1]^2 \times [0,1]$.
- *Phases.* For particles on a circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the phase comparison is the relation observable of $\{(x, y) : y - x \in (0, \frac{1}{2}) \bmod 1\}$, the event "Bob's angle is ahead of Alice's by less than half a turn".
- *Spin chains.* Alice holds a chain of spins of variable finite length, with Hilbert space $\ell^2(\{0,1\}^{<\mathbb{N}}) \cong \bigoplus_{n\ge 0}(\mathbb{C}^2)^{\otimes n}$, where $\{0,1\}^{<\mathbb{N}}$ is the set of finite sequences of zeros and ones; for $t \in \{0,1\}^{<\mathbb{N}}$, the basis ket $|t\rangle = \bigotimes_{i<|t|} |t(i)\rangle$ is the chain of length $|t|$ whose spin at site i, for $i < |t|$, is up exactly when $t(i) = 0$. In *scanning*, or reading site by site, Bob's label is a site i and the observable reads 1 when Alice's spin at i is down, the test $|1\rangle\langle 1|$ at site i, whose value on $|t\rangle$ is $t(i)$ when $i < |t|$ and 0 when the chain is shorter. In *template matching*, Bob's label is a finite reference configuration u and the observable reads 1 when Alice's chain is an initial segment of u. The *combined* observable offers Bob both kinds of label. All three are diagonal observables.
- *Fringes.* The *wavenumber* of monochromatic light is $k = 2\pi/\lambda$, with λ the wavelength. Behind two slits, the two paths to a point x of a distant screen differ in phase by kx, in suitable units, so the amplitudes add to a multiple of $1 + e^{ikx}$ and the probability density is proportional to $|1 + e^{ikx}|^2 = 2(1 + \cos kx)$. Alice's register holds the wavenumber and Bob's the detector position, and the fringe observable is $C = \sum_n |e_n\rangle\langle e_n| \otimes M_{f_n}$ on $\ell^2(\mathbb{N}) \otimes L^2[0, 2\pi]$, where $(e_n)_{n\in\mathbb{N}}$ is the standard basis of $\ell^2(\mathbb{N})$ and each M_{f_n} is multiplication by $f_n(x) = \frac{1}{2}(1 + \cos k_n x)$, for a sequence $(k_n)_{n\in\mathbb{N}}$ of wavenumbers.
- *The Ising family.* The example of Variatio 17 is a family of response functions, indexed by finite samples with offsets and by applied fields, to which the classification applies through the pointwise closure of the family.

The model-theoretic reading. Let C be a bipartite observable, and let $\mathcal{M}_C$ be its two-sorted structure (see Variatio 1). The compactum classified here is the space of idealized responses of C, that is, of the φ-types over a large elementary extension of $\mathcal{M}_C$ that are approximately finitely satisfiable in $\mathcal{M}_C$ (see Variatio 13), the coheirs of Variatio 1; by Theorem 11.1 it is a Rosenthal compactum when C is NIP. Its topology is that of pointwise convergence on the parameters, so metrizability says that the values at countably many parameters determine every such type, and first countability says that each type is determined, among all of them, by countably many parameters chosen for it. When metrizability fails, the trichotomy names the configuration of types that must be present: split pairs of types, like the two cuts at a point of the rational order, an uncountable discrete family of types, or a single type that is the limit of continuum many types isolated from one another.

For the physicist or operator theorist. The compactum classified here is the image of the state space of $B(H_1)$ under the map $\omega_1 \mapsto \mathfrak{r}_{\omega_1}$, with the topology of pointwise convergence on the probe space, and the trichotomy measures how far this image is from metrizable.

For the logician. Metrizability has an experimental meaning: there is a fixed countable list of Bob's probes whose answers determine the response of every idealized preparation of Alice to every probe. For the comparison of photon numbers such a list exists (see Theorem 16.1), because a response can jump only at the countably many points of the spectrum. For the comparison of positions no countable list suffices, since an idealized response can jump at any point of the line, but each single idealized response is determined by probes at countably many points chosen for it. For the causal order even this fails: an event whose light cone has shrunk to a point c of Bob's time slice is distinguished only by probes that assign positive probability to c itself (see Variatio 16).

Outside the metrizable class, which contains the stable observables, each class of the trichotomy is thus a degree of failure of the double limit condition that stops short of the independence property, and for observables one of these classes is empty (see Theorem 19.1).

VARIATIO 15

a 1 Clav. Canone alla Quinta. Andante

The heptachotomy

In this variation we realize each of the seven minimal families of the heptachotomy of Argyros, Dodos and Kanellopoulos by a selected tree of physical preparations: photons, particles and spins. The heptachotomy refines Todorčević's trichotomy (see Variatio 14). Where the trichotomy identifies canonical compacta, the heptachotomy records how families indexed by the dyadic tree sit inside their closures; the logician classifies the indexed families that lead to a given compactum, and the physicist identifies the compacta reached by trees of preparations. A tree of preparations branches in two at every node and gathers continuum many paths into a single countable family; in the realizations below, the responses along each branch converge, and their limits belong to the pointwise closure of the selected family. That closure need not be convex, whereas the full response compactum of the observable contains the responses of mixtures and is convex; the observables themselves are placed in their classes by Theorem 16.1 of the next variation.

The heptachotomy describes families of pairwise distinct functions indexed by the dyadic tree $\{0,1\}^{<\mathbb{N}}$ of finite sequences of zeros and ones. A *dyadic subtree* of $\{0,1\}^{<\mathbb{N}}$ is a subset that, ordered by extension, is isomorphic to $\{0,1\}^{<\mathbb{N}}$; we index it as $(s_t)_{t\in\{0,1\}^{<\mathbb{N}}}$ by its canonical enumeration, the unique such isomorphism $t \mapsto s_t$ that also preserves the lexicographic order. Two families $\{f_t : t \in \{0,1\}^{<\mathbb{N}}\}$ and $\{g_t : t \in \{0,1\}^{<\mathbb{N}}\}$ are *equivalent* if the map $f_t \mapsto g_t$ extends to a homeomorphism of their pointwise closures, and such a family that is relatively compact in the first Baire class of a Polish space, the class of pointwise limits of sequences of continuous functions on it (see Variatio 11), is *minimal* if it is equivalent to the family $\{f_{s_t} : t \in \{0,1\}^{<\mathbb{N}}\}$ for every dyadic subtree $(s_t)_{t\in\{0,1\}^{<\mathbb{N}}}$. There are exactly seven minimal families $D_1, \dots, D_7$ up to equivalence, and every family of pairwise distinct functions indexed by $\{0,1\}^{<\mathbb{N}}$ that is relatively compact in the first Baire class of a Polish space contains, along a *regular* dyadic subtree, that is, one each of whose levels lies in a single level of $\{0,1\}^{<\mathbb{N}}$, a family equivalent to one of them [**ADK08**, Theorem 2]. Their closures fall into the classes of the trichotomy (see Variatio 14): those of D_1 and D_2 are metrizable, namely the one-point compactification $A(2^{<\mathbb{N}})$ of the countable discrete set $\{0,1\}^{<\mathbb{N}}$ and the compactum $2^{\leq\mathbb{N}}$ of the finite and infinite sequences of zeros and ones, respectively; those of D_3 and D_4 are, respectively, the split Cantor sets $\hat{S}_+(2^{\mathbb{N}})$ and $\hat{S}_-(2^{\mathbb{N}})$, those of D_6 and D_7 are, respectively, the duplicates $\hat{D}(2^{\mathbb{N}})$ and $\hat{D}(S(2^{\mathbb{N}}))$, and that of D_5 is the compactum $\hat{A}(2^{\mathbb{N}})$.

Family	Closure	The picture	The physics
D_1	$A(2^{<\mathbb{N}})$	escape to infinity	photon number states
D_2	$2^{\leq\mathbb{N}}$	infinite configurations	a spin chain read site by site
D_3	$\hat{S}_+(2^{\mathbb{N}})$	approach from the right	a particle in a box
D_4	$\hat{S}_-(2^{\mathbb{N}})$	approach from the left	a particle in a box
D_6	$\hat{D}(2^{\mathbb{N}})$	twins over a Cantor set	a spin chain, sites and templates
D_7	$\hat{D}(S(2^{\mathbb{N}}))$	approach from both sides	a particle in a box, half on each side
D_5	$\hat{A}(2^{\mathbb{N}})$	a point-supported limit	light cones; templates

TABLE 5. The seven minimal families, grouped by class of Todorčević's trichotomy.

Theorem 15.1. *Each of the seven minimal families is realized by a tree of pure preparations for one of the observables of Variatio 14: the family D_1 by photon number states, compared by photon number; the family D_2 by a spin chain read site by site; the families D_3 and D_4 by a particle in a box spread over a gap of a Cantor set, compared by position; the family D_5 by a spin chain matched against templates; the family D_6 by a spin chain read both site by site and against templates; and the family D_7 by a particle in a box spread half over each of two short intervals that flank an interval of the construction of a Cantor set. The causal order of a $(1+1)$-dimensional spacetime also realizes D_5, in the sense that, for every sequence of Alice's density operators that is dense in trace norm, the responses of some subfamily, indexed by $\{0,1\}^{<\mathbb{N}}$, are equivalent to D_5.*

The realizations display what distinguishes the families (see Figure 6), and the display is sharpest in the limits of the responses along three routes towards an infinite sequence b of zeros and ones with infinitely many of each, recorded in Table 6 and drawn in Figure 5; these are the branches that admit departures on both sides. Along the branch of b, the nodes are the initial segments of b; departing below, they converge to b, are not initial segments of b, and satisfy $t^\frown 0^\infty < b$ in the lexicographic order; departing above, they satisfy $t^\frown 0^\infty > b$ instead. For the comparison of positions, x_b is the point of the Cantor set at which the supports of the preparations indexed by the branch accumulate.

Three comparisons stand out. A particle that approaches from one side follows, along the branch, the convention of the departures below (D_3) or of those above (D_4), and the side of approach is the only difference between the two families. A particle spread over both sides has, along the branch, a third limit, distinct from both departures (D_7). Template matching and the combined observable share the limit along the branch, an exact match with the template of b, and differ in the departures: in D_5 every departure ends in the same zero response, whatever b is, and this common limit has no countable base of neighborhoods; in D_6 the departures keep the configuration b, which the site probes read, and this restores first countability.

Equivalent families have homeomorphic closures, not identical functions, and the reading that follows describes the canonical families. The functions that make up the seven families are primitives of binary search on the dyadic interval of a node t, the set of infinite sequences that extend t: the prefix test v_t, which asks whether a sequence extends t; the rounded value

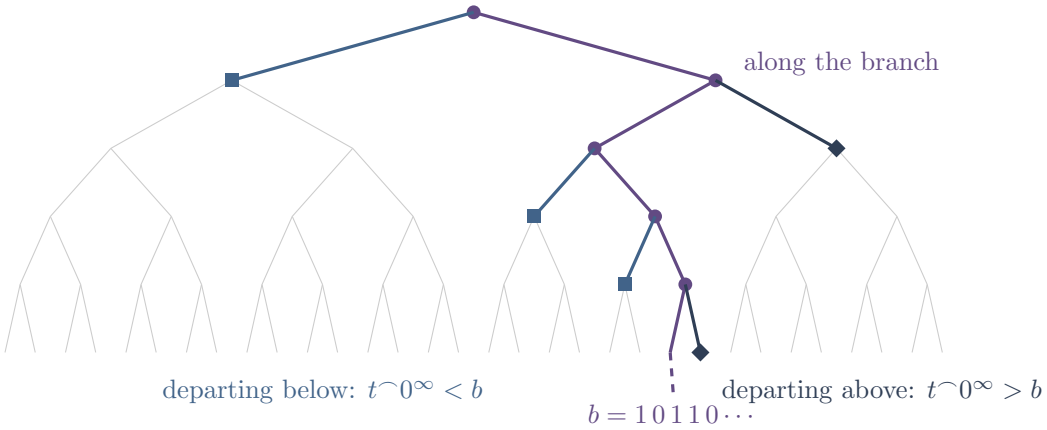


FIGURE 5. The three routes towards a branch b of the dyadic tree, ordered lexicographically from left to right.

Family	Observable	Along the branch	Departing below	Departing above
D_1	photon numbers	infinitely many photons	infinitely many photons	infinitely many photons
D_2	scanning	the configuration b	the configuration b	the configuration b
D_3	positions	just right of x_b	just right of x_b	just left of x_b
D_4	positions	just left of x_b	just right of x_b	just left of x_b
D_5	templates	matched by the template of b	no match	no match
D_6	sites and templates	b, matched by its template	b, matched by no template	b, matched by no template
D_7	positions, two halves	half on each side of x_b	just right of x_b	just left of x_b

TABLE 6. Limits of the responses along the three routes.

$t^\frown 0^\infty$; and the comparisons with the two endpoints of the interval. In the realizations these primitives become measurements, on the point probes at which the realizations reproduce the canonical functions: the infinite configurations for templates and sites, and the points of the Cantor set for positions. Template matching performs the prefix test (D_5), the spin chain read site by site reports the rounded value (D_2), and the combined observable reports both (D_6). A particle spread over the gap beside an endpoint compares Bob's particle with that endpoint (D_3, D_4), and a particle spread half over each side answers a three-way comparison, below, inside or above (D_7). Photon number states are not prefix tests, but every sequence of distinct number states has one and the same idealized limit, as every sequence of distinct weighted prefix tests converges to zero (D_1).

The families D_3 and D_4 appear although the hereditarily separable non-metrizable class is empty for observables. There is no contradiction: the closure of the family itself is a split

Cantor set, but the idealized responses of the whole observable include the mixtures, and, by Theorem 19.1, they form a larger compactum, which is not hereditarily separable.

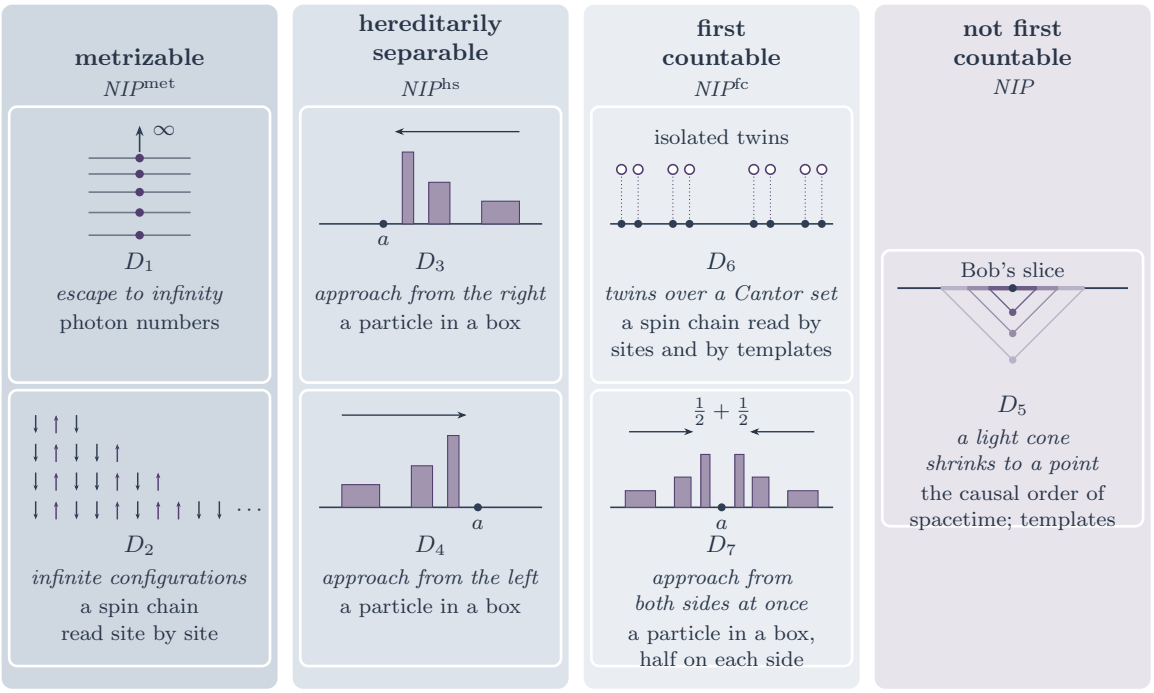


FIGURE 6. Schematic realizations of the seven minimal families of the heptachotomy, arranged by the classes of Todorčević's trichotomy.

The model-theoretic reading. A tree $(\rho_t)_{t\in\{0,1\}^{<\mathbb{N}}}$ of preparations is a tree of elements of Alice's sort in the structure $\mathcal{M}_C$ that a bipartite observable C defines. The pointwise closure of the responses of the tree is taken on Bob's probes, including idealized probes, with Bob's system idealized first (see Variatio 13). On actual probes, an ultrafilter limit along nodes $(t_n)_{n\in\mathbb{N}}$ restricts to the φ-type over $\mathcal{M}_C$ of the ultrapower element $[\rho_{t_n}]_{\mathcal{U}}$, where $\mathcal{U}$ is an ultrafilter on $\mathbb{N}$ (see Variatio 1); two idealized responses may have the same restriction while differing on idealized probes. When C is NIP and the responses of the nodes are pairwise distinct, passing to a regular dyadic subtree, by a Ramsey theorem for trees, plays the role that the extraction of indiscernible sequences plays in model theory, and the heptachotomy says that, after this thinning, the indexed family of responses is equivalent to one of the seven minimal families. The class of the trichotomy to which its closure belongs, recorded in Table 5, describes which countable collections of probes suffice to distinguish the responses.

For the physicist or operator theorist. The heptachotomy is a Ramsey theorem for trees: after passage to a regular dyadic subtree, the closure of a tree of responses is one of seven canonical compacta, up to a homeomorphism that respects the indexing.

For the logician. In each realization the nodes of the dyadic tree are pure states of Alice's system, and the idealized limits that illustrate each family are taken along the branches, without exhausting the closure. For photon numbers the nodes are states with definite, pairwise distinct numbers of photons, so every branch escapes to infinite photon number

and all branches share one idealized limit (D_1); for a spin chain read site by site the node t is the chain $|s_t\rangle$ of the regular dyadic subtree (s_t) selected in section P.2, chosen so that the scanning responses, the sequences s_t followed by zeros, are pairwise distinct; along a branch the scanning responses of the chains converge to those of the infinite configuration that extends all the s_t of the branch, whose response to the down-spin test at site i is 1 if the spin at i is down and 0 if it is up (D_2). For a particle in the box $[0, 1]$, compared by position with Bob's particle, the nodes are wave functions spread uniformly over short intervals that shrink, along each branch, to a point of a Cantor set, from one side (D_3, D_4) or half on each side (D_7); the idealized limit answers like a particle at the limit point, except on probes that assign positive mass to the point itself, where the side of approach determines the answer.

A minimal family belongs to a selected tree of preparations, and not to an observable. The comparison of positions realizes D_3, D_4 and D_7; the spin chain realizes D_2, D_5 and D_6; the causal order realizes D_5. The class of an observable is that of its whole response compactum, which contains the responses of mixtures and, by Theorem 19.1, is never hereditarily separable without being metrizable. The minimal family is decided by the selected tree and by the probes that read its responses. In the Cantor-gap trees constructed here for D_3, D_4 and D_7 the departures on the two sides of a branch end in distinct limits, a particle just left and a particle just right of a point, whereas the departures of D_6 share one limit; whether some tree of the comparison of positions realizes D_6 is open (see Question 4).

VARIATIO 16

a 1 Clav. Ouverture

Physical orders in every class

Each of the three classes available to full convex NIP response compacta contains a physical order, and the interference fringes of the double slit (see Variatio 11) lie beyond NIP. In the examples below, what separates the classes is the way in which the idealized responses depart from continuity: by jumps confined to countably many points, by jumps that may occur at any point of a continuum, or by deviations concentrated at single points. In this variation we state the result and then treat the examples in turn. In Figure 7 we draw, for parts (a), (b), (c) and (e), a monotone response on the relabeled photon-number spectrum $\{1-\frac{1}{l+1} : l \in \mathbb{N}\}$, two idealized position responses with the response of the equal mixture of the two idealized preparations, light cones of events approaching Bob's time slice, and fringes that realize the pattern $1, 2, 4, 8$ at a single detector position.

Theorem 16.1.

(a) *The comparison of two photon numbers is unstable and* $\mathrm{NIP}^{\mathrm{met}}$.

(b) *The comparison of two positions on a line, or of two quadratures of light, is* $\mathrm{NIP}^{\mathrm{fc}}$ *but not* $\mathrm{NIP}^{\mathrm{hs}}$*; its idealized responses form a copy of the space of nondecreasing functions from* $[0,1]$ *to* $[0,1]$ *with value* 0 *at* 0 *and* 1 *at* 1*, a closed subspace of the Helly space.*

(c) *The causal order of a* $(1+1)$*-dimensional spacetime, "Alice's event can influence Bob's detection", is NIP but not* $\mathrm{NIP}^{\mathrm{fc}}$.

(d) *The comparison of two phases on a circle, "Bob's angle is ahead of Alice's by less than half a turn", is NIP but not* $\mathrm{NIP}^{\mathrm{fc}}$.

(e) *The interference fringes of wavenumbers* 8^n*, for* $n \geq 1$*, are not NIP: the closure of their responses contains a homeomorphic copy of* $\beta\mathbb{N}$.

Photons. The idealized responses of a comparison are monotone, and a monotone function has at most countably many jumps, each with two sides. For discrete spectra, every idealized response jumps only at the countably many points of the spectrum, so countably many probes separate them. Every comparison is moreover uniformly NIP, in the sense that, for every $\delta > 0$, its δ-shattered sets of responses, the finite sets on which probes realize every pattern with a gap of at least δ, have bounded size, and the bounds do not depend on the Hilbert spaces: replacing the distribution of the compared quantity in Bob's state by m equally weighted quantile points changes every response by at most $1/m$, and only polynomially many patterns survive the discretization (see Proposition P.35 and Corollary P.36).

Positions. For continuous spectra every point can be a jump, approached from either side, and the idealized responses form a copy of a closed subspace of the Helly space.

Light cones. In $1+1$ dimensions, an event of Alice influences a detection of Bob exactly when the detection lies between the two points where the light rays through the event cross

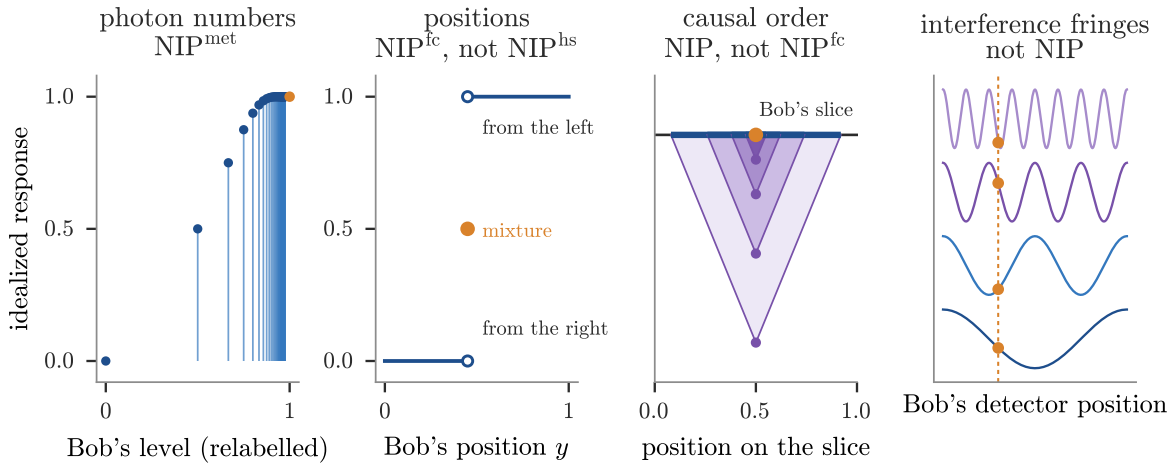


FIGURE 7. The physical orders of Theorem 16.1.

Bob's time slice. Light-cone sections that contain a fixed point and shrink to it yield an idealized response whose profile on point probes is supported at the fixed point; on an arbitrary probe measure, the response reads the mass assigned to this point. This is the picture of the compactum $\hat{A}(2^{\mathbb{N}})$ that witnesses the failure of first countability.

Phases. The responses of particles whose angle is placed half just behind a point of the circle and half just ahead of the antipodal point converge, as the two arcs shrink, to an idealized response whose profile on point probes equals $\frac{1}{2}$, the response of the uniform state, except at these two points; for a comparison on a line no response of this kind exists, since its profiles are monotone, and a monotone profile that is constant away from finitely many points is constant.

Fringes. For each $n \geq 1$, the response $\frac{1}{2}(1+\cos 8^n x)$ of a photon prepared with wavenumber 8^n to a detector at position x depends only on $8^n x$ modulo 2π, that is, on the base-8 digits of $x/2\pi$ from the $(n+1)$st on, and these digits can be chosen to realize every finite pattern.

The model-theoretic reading. Each observable of parts (a)–(d) of Theorem 16.1 reads a relation of a classical structure, whose types display the classes. In the structure of the causal order, the two sorts are $[0,1]^2$, the pairs (a,b) of points where the two light rays through an event of Alice cross Bob's time slice, and $[0,1]$, the detector positions y; the relation, which we write $\varphi((a,b),y)$, is $a \leq y \leq b$. For a point c with $0 < c < 1$ and a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the element $[(c-1/n,\, c+1/n)]_{\mathcal{U}}$ of the ultrapower, the class of a sequence indexed by the integers $n > 1/\min(c, 1-c)$, is a pair of endpoints of an interval of infinitesimal length around c, and its φ-type over the real points is $y \mapsto 1_{\{c\}}(y)$; these types, one for each c, are isolated from one another and accumulate at the zero function, which therefore has no countable base of neighborhoods, the configuration of $\hat{A}(2^{\mathbb{N}})$. For the comparison of positions the corresponding types are the split pairs of Variatio 12, and the idealized responses of the observable, which include those of mixtures, form a copy of the closed subspace of the Helly space that consists of the nondecreasing functions from $[0,1]$ to $[0,1]$ with value 0 at 0 and 1 at 1. In the quantum structure $\mathcal{M}_C$ of each such observable C (see Variatio 1), these types appear as the restrictions of idealized responses to Bob's idealized probes concentrated at points, with Bob's system idealized first (see Variatio 13).

For the physicist or operator theorist. The idealized responses of these observables are integrals of their profiles on point probes against the probe measures, and the class is read off the profiles: jumps at countably many points for photon numbers, jumps at any point for positions, and point masses for light cones.

For the logician. In a $(1+1)$-dimensional spacetime with coordinates (t, x) and light speed 1, a signal emitted at an event reaches only the events inside its future light cone, so the event (t, x) with $t \leq 0$ can affect a detector at the point y of the slice $t = 0$ exactly when $x + t \leq y \leq x - t$. The causal-order observable asks this question of an event of Alice and a detector position of Bob, both described by quantum states, and answers yes with the probability that the two positions satisfy the inequality. For events whose light-cone sections contain a fixed point c and shrink to it as the events approach the slice, the idealized limit has response $\nu(\{c\})$ at a probe represented by the probability measure ν; since no actual position distribution assigns positive probability to a single point, this idealized response is detected only by idealized probes. Each probe measure has only countably many atoms, so continuum many such point-supported profiles cannot be monitored by countably many probes. For the phases the corresponding feature is that a circle, unlike a line, carries an idealized response whose profile on point probes equals $\frac{1}{2}$ except at two antipodal points.

Every order considered here is unstable, and the topology of the response compactum tells the instabilities apart: jumps at countably many points, jumps at any point of a continuum, deviations concentrated at single points. These realizations have commutative probe algebras, so their responses can be described by integration against probability distributions; the classification itself applies to every bounded self-adjoint bipartite observable, without the commutativity assumption. The thermodynamic families of the next two variations apply the same topology of pointwise closures to a different question, the order of limits in a phase transition.

VARIATIO 17

a 2 Clav.

Orders of limits in phase transitions

In this variation we place spontaneous magnetization, the standard instance of two limits that cannot be interchanged, in the classification of response families. The magnetization of a ferromagnet in a field, taken to infinite volume and then to zero field, is nonzero below the critical temperature, whereas the other order of the limits yields zero. The distinction between stability and instability already appears without translating the field; offsets that move the transition point reveal the finer topological class.

The Ising model is the basic model of a ferromagnet. Fix a dimension $d \geq 1$. On the box $\Lambda_N = \{-N, \dots, N\}^d$ of the lattice $\mathbb{Z}^d$, where N is its half-width, with periodic boundary conditions, a configuration $s \in \{-1, 1\}^{\Lambda_N}$ has energy $E(s) = -J \sum_{i \sim j} s_i s_j$, where the sum runs over the pairs of neighboring sites and the coupling J is positive. At inverse temperature $\beta > 0$ and applied field h, measured as an energy, the *Gibbs state* assigns to s the probability $Z_N^{-1} e^{-\beta(E(s) - h \sum_{i \in \Lambda_N} s_i)}$. The normalizing constant $Z_N = \sum_s e^{-\beta(E(s) - h \sum_{i \in \Lambda_N} s_i)}$ is the *partition function*; the magnetization $m_N(h)$ is the expected mean spin in the Gibbs state.

The *spontaneous magnetization* is $m^* = \lim_{h \downarrow 0} \lim_N m_N(h)$, along boxes exhausting $\mathbb{Z}^d$; for $d \geq 2$ there is a critical β_c with $m^* > 0$ for $\beta > \beta_c$, the *ordered phase*, and $m^* = 0$ for $\beta < \beta_c$. The *pure phases* are the limits, as $h \downarrow 0$ and as $h \uparrow 0$, of the infinite-volume limits of the Gibbs states at field h, with mean spin $\pm m^*$.

For the classification we use a parameterized response family. The preparations are finite samples indexed by N and field offsets $a \in [-1/\beta, 1/\beta]$, and the probes are applied fields $h \in [-2/\beta, 2/\beta]$; the response is the magnetization $m_N(h - a)$. The probes are values of an applied-field parameter, not preparations of a second system. The family has two variables, and its classification does not require identifying it with the full response family of a single bounded bipartite observable. In the ordered phase the offset a moves the jump of the infinite-volume magnetization from zero field to the field $h = a$; the offsets thus supply a continuum of jumps, one for each offset, and it is this continuum that the classes of Todorčević's trichotomy detect.

For the Ising model in dimension $d \geq 2$, the family with offsets is stable exactly when the spontaneous magnetization vanishes. When the spontaneous magnetization is positive, this family is $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$, and the idealized responses of the family form a first countable compactum that is not hereditarily separable: the plus and minus phases yield two idealized responses that differ at a single point, a *split pair*, and the midpoints of these pairs, one for each offset a, form an uncountable discrete set (see Theorem P.54). The family $\{m_N\}$ without offsets is also stable exactly when $m^* = 0$, but its closure is metrizable in both regimes (see Theorem P.54). The split pair is illustrated in Figure 8 in the mean-field (Curie–Weiss) model, where the magnetization of N spins can be computed exactly; in the

ordered phase the curves steepen into a jump at $h = 0$ between the values $\pm m^*$, and in the disordered phase they stay smooth.

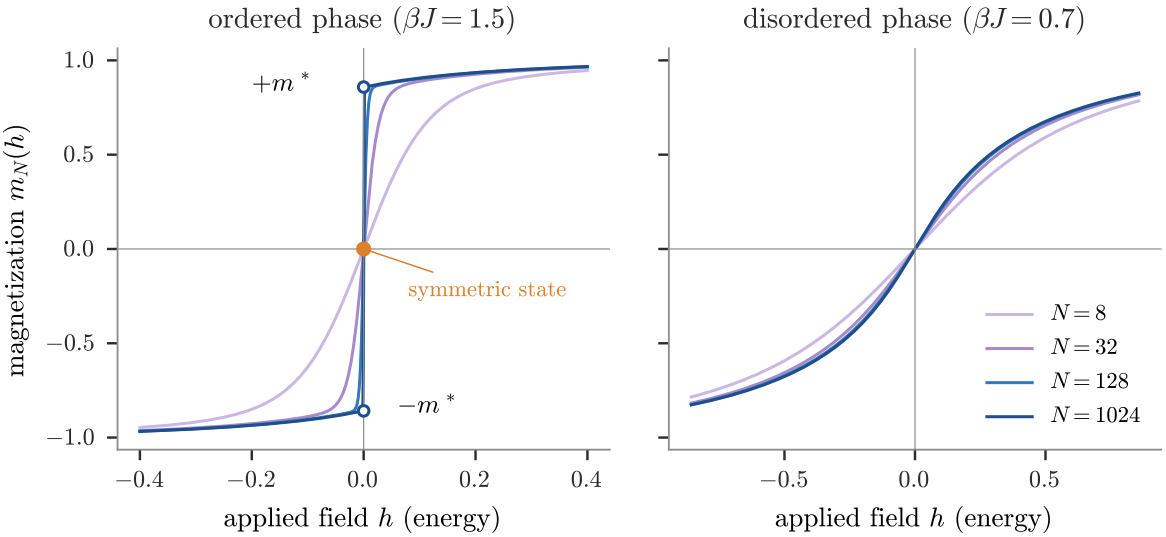


FIGURE 8. Magnetization curves of the Curie–Weiss model for increasing numbers N of spins, in the ordered and in the disordered phase.

The model-theoretic reading. The magnetizations $m_N(h - a)$ form a formula in the sense of Variatio 1, with finite samples, indexed by size and offset, and applied fields as the two sorts, and the idealized responses are its φ-types realized in ultrapowers. In the ordered phase, two of them, the plus and the minus phase at offset a, take the values m^* and $-m^*$ at the parameter $h = a$ and agree at every other parameter; they form a split pair, like the types just above and just below a real number in the rational order (see Variatio 12). Each response is monotone in h, so the family is NIP, and the types that take the midpoint value 0 at $h = a$, one for each offset a, form an uncountable discrete set, which places the family in the class $\mathrm{NIP}^{\mathrm{fc}} \backslash \mathrm{NIP}^{\mathrm{hs}}$; in the disordered phase the idealized responses are continuous and the family is stable.

For the physicist or operator theorist. The two pure phases at an offset are the one-sided limits, at the transition field, of the infinite-volume magnetization curve, and the trichotomy classifies the family of these curves.

For the logician. The Ising model describes a magnet by a spin $s_i = \pm 1$ at each site of a lattice, and thermal equilibrium at inverse temperature β by the Gibbs state; the magnetization $m_N(h)$ is the expected mean spin, an odd nondecreasing smooth function of the applied field h for each finite N. The two iterated limits in question are $\lim_{h \downarrow 0} \lim_N m_N(h) = m^*$ and $\lim_N \lim_{h \downarrow 0} m_N(h) = 0$: a finite magnet in zero field has no preferred direction, while an infinite magnet in a vanishing positive field keeps a positive magnetization. In the bipartite family, a preparation is a finite sample whose field is shifted by the offset a and a probe is the applied field. An idealized preparation is represented by an infinite-volume limit. For an interior offset $a \in (-1/\beta, 1/\beta)$ its response at the field $h = a$ may take any value between $-m^*$ and m^*. The values m^* and $-m^*$ arise when the field reaches the transition from one

side after the volume has become infinite. These two extreme responses are the two pure phases, and their equal mixture, whose mean spin is 0, is the midpoint.

The trichotomy thus does more than detect a phase transition: the ordered phase places the magnetization in the class $\mathrm{NIP}^{\mathrm{fc}} \backslash \mathrm{NIP}^{\mathrm{hs}}$, through the split pair of the two pure phases, and the disordered phase leaves it stable.

VARIATIO 18

a 1 Clav. Canone alla Sesta

A quantum ferromagnet in the classification

A magnetic field perpendicular to the axis along which the spins of an Ising magnet align makes the model quantum. In this variation we carry the results of the previous one over to the quantum model.

In a transverse field the Ising Hamiltonian on Λ_N (see Variatio 17) is $H_N = -J\sum_{i\sim j}\sigma_i^z\sigma_j^z - \Gamma\sum_{i\in\Lambda_N}\sigma_i^x$, where each site i carries a qubit, σ_i^z and σ_i^x are the Pauli matrices $\mathrm{diag}(1,-1)$ and $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ acting on the qubit at i (see Variatio 22), the first sum runs over the pairs of neighboring sites, the coupling J is positive, and the transverse-field strength Γ is nonnegative. For $\Gamma = 0$ the Hamiltonian is diagonal in the basis of configurations and is the classical Ising energy at zero field; for $\Gamma > 0$ the Hamiltonian is not diagonal and does not commute with the magnetization, and the Gibbs state at zero field is the density operator $e^{-\beta H_N}/\operatorname{Tr} e^{-\beta H_N}$. Write $S_N = \sum_{i\in\Lambda_N}\sigma_i^z$ for the total spin along the z axis. At inverse temperature β (see Variatio 17), the *partition function* in a field h, real or complex, is $Z_N(h) = \operatorname{Tr}\exp(-\beta(H_N - hS_N))$, and the finite-volume pressure is $p_N(h) = |\Lambda_N|^{-1}\log Z_N(h)$, that is, $-\beta$ times the free energy per site. The infinite-volume pressure is $p(h) = \lim_N p_N(h)$ for real h; it is convex, and where it is differentiable the infinite-volume magnetization is $p'(h)/\beta$. The *magnetization with offsets* of the quantum model is the family of responses $m_N(h-a)$ of Variatio 17, now computed in the Gibbs state of H_N.

The classification of Variatio 17 uses the differentiability of the pressure away from zero field. For classical ferromagnets this follows from the Lee–Yang theorem [**LY52**]: for a classical Ising ferromagnet on a finite set, the partition function, as a function of a complex field h, vanishes only where $\operatorname{Re} h = 0$. The theorem below gives the same zero-free half-planes in a transverse field.

Theorem 18.1. *For every N, the function Z_N satisfies $Z_N(h) \neq 0$ whenever $\operatorname{Re} h \neq 0$; moreover, the infinite-volume pressure $\lim_N |\Lambda_N|^{-1}\log Z_N(h)$ is real analytic for real $h \neq 0$, and the magnetization with offsets is stable when the spontaneous magnetization vanishes and NIP$^{\mathrm{fc}}$ but not NIP$^{\mathrm{hs}}$ when it is positive.*

The idea is to approximate $\exp(-\beta(H_N - hS_N))$ by the Lie–Trotter formula, which turns the quantum model into a classical ferromagnet in one more dimension, to apply the classical Lee–Yang theorem to the approximants, and to pass the zero-free half-planes to the limit by Hurwitz's theorem. The proof is a direct derivation of the analytic input that the classification needs (see section P.2 and the historical remarks).

The Lie–Trotter formula is $e^{A+B} = \lim_n (e^{A/n}e^{B/n})^n$ for square matrices A and B of the same size, where n runs through the positive integers. Applied to $-\beta(H_N - hS_N)$, it writes $Z_N(h)$ as a limit of sums over configurations of n stacked copies of the lattice; for $\Gamma > 0$, each factor $e^{\beta\Gamma\sigma_i^x/n}$ couples neighboring copies ferromagnetically, because, with $\varepsilon = \beta\Gamma/n$ and $|s\rangle$ the eigenvector of σ^z with eigenvalue $s \in \{-1, 1\}$, the matrix entries are $\langle s|e^{\varepsilon\sigma^x}|s'\rangle = c\,e^{K_\varepsilon ss'}$ for $K_\varepsilon = -\frac{1}{2}\log\tanh\varepsilon > 0$ and $c = (\sinh\varepsilon\cosh\varepsilon)^{1/2}$. Each slice has spatial coupling $\beta J/n$

and field coefficient $\beta h/n$. The result is a classical Ising ferromagnet on the lattice times a cycle of length n, to which the Lee–Yang theorem applies. For $\Gamma = 0$ the partition function is already classical, so no interslice coupling formula is needed.

The theorem is illustrated in Figure 9 on a ring of six spins, with $\beta = J = 1$. For $\Gamma = 1$ (left), the zeros, computed by exact diagonalization, appear as the dark spots of the shading $\log\big(|Z(h)|/Z(\operatorname{Re} h)\big)$, and they remain on the imaginary axis as the transverse field grows from $\Gamma = 0$ (right).

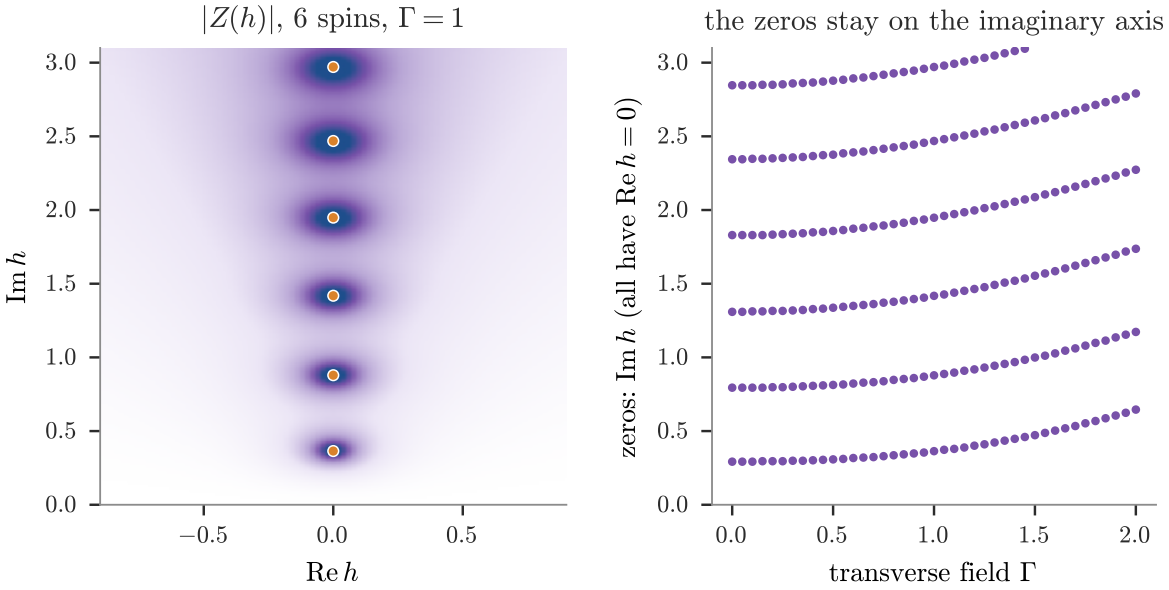


FIGURE 9. Zeros of the partition function of the Ising model in a transverse field on a ring of six spins.

The model-theoretic reading. The formula being classified is the family $h \mapsto m_N(h-a)$. Its preparation labels are finite samples indexed by N and field offset a; its probe parameter is the applied field h (see Variatio 17). Each response is nondecreasing in h; at fixed thresholds $u < v$, two responses cannot realize both a low/high pattern and its high/low reversal, so the family is NIP. The finer classification uses the closure, and the analyticity of the pressure away from zero field supplies the regularity of the limiting magnetization (see Theorem 18.1). When the spontaneous magnetization is positive, translating the field supplies split pairs at a continuum of transition points, and their midpoints form the uncountable discrete set that excludes $\mathrm{NIP}^{\mathrm{hs}}$; the closure is first countable, so the family is $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$. When the spontaneous magnetization vanishes, all idealized responses are continuous and the family is stable. Without offsets, the family is $\mathrm{NIP}^{\mathrm{met}}$ in both cases (see Theorem P.58).

For the physicist or operator theorist. The result of Theorem 18.1 concerns the Gibbs states of a Hamiltonian that does not commute with the magnetization: although $\exp(-\beta(H_N - hS_N))$ is not diagonal in the basis of configurations, its trace has the zero-free half-planes of a classical ferromagnet, and the magnetization, the expectation of $S_N/|\Lambda_N|$ in these states, falls into the classes of Variatio 17.

For the logician. In the transverse-field model each site carries a qubit, the term $-J\sigma_i^z\sigma_j^z$ favors neighbors aligned along the z axis, and the term $-\Gamma\sigma_i^x$ favors at each site the equal superposition $|+\rangle$ of the spin-up state $|0\rangle$ and the spin-down state $|1\rangle$, a state that is not an eigenvector of σ_i^z; when $\Gamma > 0$, the two terms do not commute and the Gibbs density operator is not diagonal in the basis of spin configurations. The magnetization is $(\beta|\Lambda_N|)^{-1}\partial_h \log Z_N(h)$, the expected value of $S_N/|\Lambda_N|$ in the density operator $\exp(-\beta(H_N - hS_N))/Z_N(h)$, and in infinite volume it is $p'(h)/\beta$ wherever the derivative exists. A phase transition occurs at a field h at which the pressure fails to be analytic. By Theorem 18.1, no such field is nonzero, so the magnetization can jump only at zero field, where the two pure phases meet.

When the spontaneous magnetization is positive, the quantum magnet fails the double limit condition exactly as the classical one does, and for the same reason: its convex pressure, real analytic away from zero field, has at zero field one-sided derivatives $\pm\beta m^*$ that differ.

VARIATIO 19

a 1 Clav.

Convexity forces metrizability

In this variation we show that mixing removes a class of Todorčević's trichotomy. By Theorem 19.1 below, for NIP observables, hereditary separability of the idealized responses implies metrizability, and only three classes therefore occur for observables: the metrizable compacta and two of the three classes of the trichotomy. The removal is the second of the three effects of mixing recorded in Table 1, and the split pairs of the comparison of positions show how the removal comes about (see Variatio 13): a split pair consists of two idealized responses whose profiles on the probes concentrated at single points differ only at one point of the line, and the equal mixture of the two idealized preparations has a third response, the midpoint of the pair.

Theorem 19.1 (first statement: Todorčević [Tod99], Gruenhage [Gru88]). *A separable convex Rosenthal compactum that is hereditarily separable is metrizable. Consequently, an NIP observable is NIP$^{\mathrm{met}}$, or NIP$^{\mathrm{fc}}$ but not NIP$^{\mathrm{hs}}$, or not NIP$^{\mathrm{fc}}$. For every sequence of Alice's density operators that is dense in trace norm, some subfamily of their responses, indexed by $\{0,1\}^{<\mathbb{N}}$, is equivalent in the second case to one of the minimal families D_6 and D_7 of the heptachotomy (see Variatio 15), and in the third to the minimal family D_5.*

The first statement holds for all locally connected Rosenthal compacta, by Todorčević's trichotomy and an argument of Gruenhage; the consequences for observables, and a short proof of the first statement for convex compacta, are in section P.2. Convex compacta are locally connected, since in the topology of pointwise convergence the relatively open convex sets form a base. For the comparison of positions, the compactum of idealized responses that contains these midpoints is a copy of a closed subspace of the Helly space (see Theorem 16.1).

The model-theoretic reading. In the structure $\mathcal{M}_C$ of a bipartite observable C (see Variatio 1), Alice's sort $\mathfrak{D}(H_1)$ is convex and the expectation φ_C is affine in Alice's variable, so every ultrapower of $\mathcal{M}_C$ along a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$ contains, with any elements $[\rho_n]_{\mathcal{U}}$ and $[\rho'_n]_{\mathcal{U}}$, the element $[\frac{1}{2}(\rho_n+\rho'_n)]_{\mathcal{U}}$, whose φ-type over $\mathcal{M}_C$ is the average of theirs; in the same way, the responses $\mathfrak{r}_\rho$ of Alice's preparations form a convex set of functions on the probe space Y_C of Variatio 11, and the idealized responses, its pointwise closure, therefore form a convex compact set. For a general NIP formula no such convexity is available, and the formula $x<y$, with parameters from the real line, has the reduced split interval as space of φ-types (see Variatio 12). Convexity excludes this class (see Theorem 19.1), and the split interval shows why, once types are replaced by probability measures on them: the midpoints of the point masses at the two types of each split pair are measures, not types of $x<y$,

each isolated among them by the indicator of the cut at its point, and they form an uncountable discrete set. In particular, the space of probability measures on the split interval is not hereditarily separable, contrary to a remark of Godefroy [**God80**].

For the physicist or operator theorist. The idealized responses are the image of the state space of $B(H_1)$ under an affine map, and therefore form a convex compact set; the theorem says that such a set cannot be hereditarily separable without being metrizable.

For the logician. Every preparation procedure can be randomized: choosing with probability $\frac{1}{2}$ between two procedures and forgetting the choice produces the mixture $\frac{1}{2}(\rho + \rho')$, whose expectations are the averages of those of ρ and ρ'. For the comparison of positions, the randomization of the two one-sided idealizations of a particle at a point r is the midpoint of the split pair at r (see Variatio 13), so mixing alone places the uncountable discrete set of midpoints among the idealized responses. Hence no quantum observable has the reduced split interval as space of idealized responses, and the comparison of positions, the closest physical counterpart of the rational order, is $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$ (see Theorem 16.1).

In Part I mixtures stood in for idealized preparations. Here they add idealized responses of their own, the midpoints, and the midpoints decide the class.

Part III. Simplicity

VARIATIO 20

a 2 Clav.

Simplicity

In this variation we state the theorem of Kim and Pillay, which characterizes simplicity, the third of Shelah's dividing lines, by axioms for a notion of independence of types. Each dividing line forbids a combinatorial configuration. Stability forbids the order property and NIP forbids the independence property; simplicity forbids a tree of instances of a formula in which every branch is consistent while, at every node, the infinitely many immediate successors are inconsistent in groups of a fixed finite size. Of the three prohibitions, only this one is compatible with the independence property: a simple theory may have it, as the theory of the random graph does (see Theorem 22.1 and Variatio 21), and still possess a notion of independence of types governed by axioms close to those of stability. Simplicity is first studied in two first-order structures, the random graph and the Pauli group; observables return in Variatio 28, where simplicity is studied for real-valued structures in the sense of Keisler [**Kei23**].

Let $\mathbb{M}$ be a monster model of a complete first-order theory T (see Variatio 1), and let $\varphi(x, y)$ be a formula. For an integer $k \geq 2$, a sequence $(b_j)_{j\in\mathbb{N}}$ is *k-inconsistent* for φ if no k of the formulas $\varphi(x, b_j)$ have a common solution. The formula has the *tree property* if there are an integer $k \geq 2$ and a tree $(b_s)_{s\in\mathbb{N}^{<\mathbb{N}}}$, indexed by the finite sequences of natural numbers, whose branches are consistent and in which, at every node, the immediate successors of the node form a k-inconsistent sequence. The formula has the *tree property of the second kind* if there are an integer $k \geq 2$ and an array $(b_{ij})_{i,j\in\mathbb{N}}$ whose rows are all k-inconsistent and in which every choice of one entry from each row is consistent. A theory is *simple* if no formula has the tree property [**She80**]. Stable theories are simple, and a simple theory without the independence property is stable [**She90**, **She80**]; the simple unstable theories are therefore exactly the simple theories with the independence property.

Simplicity has a structural characterization, due to Kim and Pillay, which is the usual means of proving that a theory is simple. Suppose that a ternary relation $A \mathop{\smile\!\!\!\!\!\!|}_C B$ on small subsets of $\mathbb{M}$ is invariant under automorphisms, symmetric, monotone ($A \mathop{\smile\!\!\!\!\!\!|}_C BD$ implies $A \mathop{\smile\!\!\!\!\!\!|}_C B$, where BD denotes $B \cup D$) and transitive ($A \mathop{\smile\!\!\!\!\!\!|}_C BD$ holds if and only if $A \mathop{\smile\!\!\!\!\!\!|}_C B$ and $A \mathop{\smile\!\!\!\!\!\!|}_{CB} D$ hold), has finite character ($A \mathop{\smile\!\!\!\!\!\!|}_C B$ if and only if $a \mathop{\smile\!\!\!\!\!\!|}_C b$ for all finite $a \subseteq A$ and $b \subseteq B$) and local character (for each finite a and each B there is $B_0 \subseteq B$ with $|B_0| \leq |T|$, the cardinality of the language of T plus $\aleph_0$, and $a \mathop{\smile\!\!\!\!\!\!|}_{B_0} B$), and satisfies extension (for all a and $C \subseteq B$ there is a' with $\mathrm{tp}(a'/C) = \mathrm{tp}(a/C)$ and $a' \mathop{\smile\!\!\!\!\!\!|}_C B$). Suppose also that the relation satisfies the *independence theorem over models*: if M is a model, $B \mathop{\smile\!\!\!\!\!\!|}_M D$, and two realizations a_1 and a_2 of the same type over M satisfy $a_1 \mathop{\smile\!\!\!\!\!\!|}_M B$ and $a_2 \mathop{\smile\!\!\!\!\!\!|}_M D$, then some a realizes $\mathrm{tp}(a_1/MB) \cup \mathrm{tp}(a_2/MD)$ with $a \mathop{\smile\!\!\!\!\!\!|}_M BD$. Kim and Pillay proved that the theory T is then simple and that the relation is nonforking independence [**KP97**]: $A \mathop{\smile\!\!\!\!\!\!|}_C B$ holds exactly when $\mathrm{tp}(A/CB)$ does not fork over C, in the sense of Variatio 1.

Stable theories have a characterization of the same kind, due to Harnik and Harrington [**HH84**], with *stationarity over models*, the uniqueness of nonforking extensions of a type over a model, in place of the independence theorem; in a stable theory, stationarity implies the independence theorem. Uniqueness is thus what separates stability from simplicity. Where uniqueness is lost, nonforking extensions of one type, chosen separately over two sets independent over a model, might conflict, and the independence theorem, the substantive axiom, ensures that they do not. The iterated ultrapower of Variatio 1 shows where nonforking amalgams come from: over a model, a coheir does not fork, so, in either order of realization, the element realized second is independent over the model of the element realized first; in a stable theory the two orders produce the only nonforking amalgam (see Variatio 6), and in a simple unstable theory other nonforking amalgams exist, whose combination the independence theorem governs. The tree property is what defeats local character: a tree with consistent branches and k-inconsistent siblings yields a type that forks over every subset of size at most $|T|$ of the domain of the type.

For the physicist or operator theorist. Simplicity concerns the combination of partial information: in a simple theory, two descriptions of an unknown element that are nonforking extensions of one type over a model, given over two sets independent over the model, can always be satisfied together, and the tree property is the configuration that prevents it.

For the logician. The structures of Part III come from the measurement of Pauli strings (see Variatio 22), tensor products of the matrices I, X, Y and Z on finitely many qubits, which are self-adjoint unitaries and therefore have no eigenvalues other than 1 and -1; each Pauli string is thus a measurement with outcomes in $\{1, -1\}$. Two observables can be measured jointly, with a joint probability distribution of outcomes in every state, exactly when they commute, and any two Pauli strings either commute or anticommute; the relation between them that matters for measurement is anticommutation. In quantum error correction an unknown error, itself a Pauli string, is probed by measuring commuting Pauli strings, each outcome recording whether the error anticommutes with the string measured, and the question that the independence theorem answers, whether partial information about an unknown element can be combined, becomes a question about these outcomes.

Simplicity is not a property of the closure of a family of responses, as stability and NIP are. Its characterization is an amalgamation property, and response closures return only when preparations are mixed: a diagonal observable that is simple, in the sense defined for its mixtures in Variatio 28, is NIP (see Theorem 28.1).

VARIATIO 21

Canone alla Settima

The Rado graph

In this variation we show that the theory of the random graph is simple and unstable, has the independence property, and has forking given by intersection. If each pair of points of a countable set is joined by an edge independently with probability $\frac{1}{2}$, the resulting graph is, with probability one, isomorphic to a fixed graph, by a theorem of Erdős and Rényi [**ER63**]. It is the first of the two basic structures of this part, and a motivating example for Shelah's simple theories.

The finite graphs form a Fraïssé class (see Variatio 4), a class of finite structures closed under substructures in which any two members can be glued along a common substructure, since two graphs that share an induced subgraph can be glued along it without new edges. The Fraïssé limit of this class is the Rado graph R, with vertex set R and edge relation E, the unique countable graph, up to isomorphism, with the *extension property*: for all finite disjoint sets P and Q of vertices, some vertex outside $P \cup Q$ is adjacent to every vertex of P and to no vertex of Q. It is the random graph of the preceding paragraph. Its theory is denoted T_{ind} [**She80**]. The relation defined, for small sets A, B and C of vertices of a monster model of T_{ind} (see Variatio 20), by letting $A \mathop{\smile\!\!\!\!\!\!|}_C B$ hold exactly when $A \cap B \subseteq C$, satisfies the Kim–Pillay axioms (see Theorem 22.1). We compute forking and verify the independence theorem, the substantive axiom, by hand, since they show how little a vertex knows.

Forking. Let $C \subseteq B$ be small sets of vertices of a monster model of T_{ind}, and let v be a vertex. The idea is that, when $v \notin B$, the type of v over B records only adjacency and distinctness, which the extension property allows one to prescribe freely, whereas, when $v \in B \setminus C$, the type names a vertex that C does not determine. If $v \in C$, the type of v over B is realized in C and does not fork over C. If $v \notin B$, the type of v over B says only which vertices of B are adjacent to v, and that v differs from each of them; it extends to the global type that, in addition, makes its variable x different from, and not adjacent to, every vertex outside B. No formula of this global type divides over C: such a formula prescribes inequalities, edges and non-edges between x and a finite tuple of distinct vertices, and, along a sequence of copies of this tuple indiscernible over C, a vertex that occurs in two copies occupies the same place in both, so the prescriptions of finitely many copies never conflict, and the extension property satisfies them by one new vertex. A complete type no formula of which divides over C does not fork over C, since it contains one of the formulas of every finite disjunction that it implies; hence the type of v over B, which the global type extends, does not fork over C either. If $v \in B \setminus C$, the formula $x = v$ divides over C, in the sense recalled in Variatio 6: the type of v over C has infinitely many realizations, so there is a sequence $v = v_0, v_1, v_2, \ldots$ of distinct vertices, indiscernible over C, and no vertex equals two of them. Hence the type $\mathrm{tp}(v/B)$ forks over C exactly when $v \in B \setminus C$.

The independence theorem. The idea is that two types over sets independent over a model prescribe edges that can meet only over the model, where the types agree. Let M be a model,

let B and D be sets with $B \cap D \subseteq M$, and let a_1 and a_2 be finite tuples of vertices that realize the same type over M, with $a_1 \mathop{\smile\!\!\!\!\!\!|}_M B$ and $a_2 \mathop{\smile\!\!\!\!\!\!|}_M D$. The two tuples have the same entries in M, at the same places, and the same pattern of equalities among their other entries; after the same rearrangement of places, we write $a_\ell = (m, a'_\ell)$, for $\ell = 1, 2$, where m lists the entries in M. The hypotheses say that a'_1 has no entry in B and a'_2 none in D, and the two types prescribe, for new vertices, one for each distinct entry of a'_1, the equalities and edges among them and to M that a'_1 and a'_2 share, the edges to $B \setminus M$ that a'_1 has, and the edges to $D \setminus M$ that a'_2 has. The prescriptions agree on M, because a_1 and a_2 have the same type over M, and they concern disjoint sets outside M, because $B \cap D \subseteq M$. Hence their union is consistent, by the extension property, and the saturation of the monster model realizes it by a tuple x with no entry in $M \cup B \cup D$. Since the type of a tuple over a set is determined by the equalities and edges within the tuple and between the tuple and the set, the tuple $a = (m, x)$, with its places restored, realizes $\mathrm{tp}(a_1/MB) \cup \mathrm{tp}(a_2/MD)$, and it satisfies $a \mathop{\smile\!\!\!\!\!\!|}_M BD$, because its entries in BD lie in M.

The extension property also yields the instability of the theory. If the vertices b_1, a_1, b_2, a_2, … are chosen in this order by the extension property, one can arrange that a_i is adjacent to b_j exactly when $i < j$: a ladder, so the theory is unstable. Similarly, for distinct vertices $a_1, \ldots, a_n$ and every set $S \subseteq \{1, \ldots, n\}$, the extension property supplies a vertex adjacent to a_i exactly when $i \in S$: the independence property.

For the physicist or operator theorist. The Rado graph is the relation in which every finite pattern of adjacency occurs, and still two descriptions of a vertex outside both parameter sets never conflict when the two sets share only a model and the descriptions agree over the model, because they prescribe edges to disjoint sets of vertices outside the model.

For the logician. The Rado graph enters measurement through its diagonal observable $C_E = \sum_{(a,b)\in E} |a\rangle\langle a| \otimes |b\rangle\langle b|$ on $\ell^2(R) \otimes \ell^2(R)$ (see Variatio 14), where $|a\rangle$ denotes the standard basis vector of $\ell^2(R)$ indexed by the vertex a: Alice's register holds a vertex, Bob's register holds a vertex, and the measurement answers yes when the two are adjacent. Since C_E is diagonal, a general preparation ρ of Alice answers every probe as the probability distribution $a \mapsto \langle a|\rho|a\rangle$ on vertices does, so superpositions of vertices behave as random vertices. The extension property says that, for any two disjoint finite sets of Alice's vertices, some actual preparation of Bob answers yes to every vertex of the first and no to every vertex of the second, and in Variatio 28 we determine what the random vertices of Alice do to the simplicity of this measurement.

Every finite pattern occurs in the random graph, so the double limit condition fails for its edge relation; nevertheless, independent information over a model can be amalgamated, and the type of a vertex forks only when the vertex itself is named.

VARIATIO 22

a 1 Clav. Alla breve

The Pauli group

In this variation we show that the Pauli group modulo phases, with its commutation relation, is a Fraïssé limit whose theory is simple and unstable. The Pauli group is the second basic structure of this part, and comes from quantum computation: the commutation of Pauli strings decides which of them can be measured jointly, and it is all that the checks of an error-correcting code record about an unknown error (see Variatio 20).

To make this precise, we start from a single quantum bit. A *qubit* is a quantum system described by the Hilbert space $\mathbb{C}^2$ (see Variatio 4), on which the Pauli matrices act:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad Y = iXZ.$$

A *Pauli string* on n qubits is a tensor product of n factors, each equal to I, X, Y or Z, such as $XZZXI = X \otimes Z \otimes Z \otimes X \otimes I$. Pauli strings are the basic measurements of quantum information, and they are labeled by vectors of $\mathbb{F}_2^{2n}$, where $\mathbb{F}_2$ is the field with two elements: for $x = (x_1, \dots, x_n)$ and $z = (z_1, \dots, z_n)$ in $\mathbb{F}_2^n$, the string with label (x, z) is the one whose k-th factor, for $k = 1, \dots, n$, is I, X, Z or Y according as (x_k, z_k) is $(0,0)$, $(1,0)$, $(0,1)$ or $(1,1)$, and we write P_v for the Pauli string with label v. Up to a phase, the product $P_v P_w$ is the string P_{v+w}; hence the Pauli strings, multiplied by the phases ± 1 and $\pm i$, form a group, the *Pauli group* on n qubits, which modulo phases is the additive group of $\mathbb{F}_2^{2n}$. The crucial fact is that two Pauli strings either commute or anticommute,

$$P_v P_w = (-1)^{\omega(v,w)} P_w P_v, \qquad \omega\big((x,z),(x',z')\big) = x \cdot z' + z \cdot x',$$

where $x \cdot z'$ is the sum $x_1 z'_1 + \cdots + x_n z'_n$ in $\mathbb{F}_2$, and that ω is a nondegenerate alternating bilinear form on $\mathbb{F}_2^{2n}$. On countably many qubits, we let $V = \bigoplus_{k \in \mathbb{N}} \mathbb{F}_2^2$, with the same form; its elements label the Pauli strings that act nontrivially on finitely many qubits. We also call (V, ω) the Pauli group, identifying each Pauli string of finite support, up to a phase, with its label.

In Figure 10 we draw the commutation pattern of the fifteen nontrivial Pauli strings on two qubits. Three pairwise commuting strings whose labels add up to zero form a triangle; there are fifteen of them, and any two commuting strings lie on exactly one. This configuration is the generalized quadrangle of order two, known in quantum information as the *doily*. In the figure, colors mark the three orbits of triangles under a Clifford unitary of order five, up to a phase.

Stabilizer codes in brief

Most entries of the dictionary of simplicity (see Table 7) are terms of quantum error correction; we summarize here what the logician needs of the subject. A quantum computer stores information in qubits; in the standard error model, errors are Pauli strings, and a code that corrects a set of Pauli errors corrects every error that is a complex linear combination

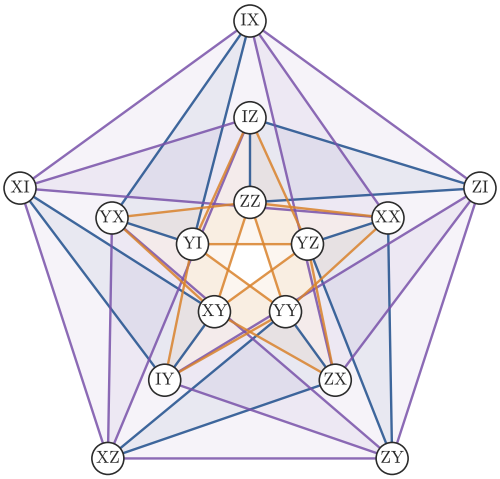


FIGURE 10. The fifteen nontrivial Pauli strings on two qubits, two of them on a common triangle exactly when they commute.

of them [Got97]. A *stabilizer group* on n qubits is a subgroup $\mathcal{G}$ of the Pauli group whose elements commute pairwise and which 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 stores k qubits [Got97]. In labels, the group $\mathcal{G}$ corresponds, up to signs, to an isotropic subspace $S \subseteq \mathbb{F}_2^{2n}$, a subspace on which ω vanishes. An error is a Pauli string e, and a Pauli string that is measured in order to detect errors is a *check*; for two Pauli strings g and e we write $\omega(g,e)$ for the value of ω at their labels. Measuring independent generators $g_1,\dots,g_{n-k}$ of $\mathcal{G}$, the checks of the code, on the corrupted state reveals the *syndrome* $(\omega(g_1,e),\dots,\omega(g_{n-k},e))$ without disturbing the stored information: if ψ lies in the code space, then we have $g_i(e\psi) = (-1)^{\omega(g_i,e)}e\psi$, so measuring g_i, that is, the projections $\frac{1}{2}(I \pm g_i)$, yields $(-1)^{\omega(g_i,e)}$ with certainty and leaves $e\psi$ unchanged, and a classical procedure, the *decoder*, infers a correction from the syndrome. An error with zero syndrome is undetectable. It is harmless if it lies in $\mathcal{G}$ up to a phase, since it then acts on the code space as a scalar, and it is a *logical error* otherwise. A *logical operator* is a Pauli string that commutes with every element of $\mathcal{G}$ without lying in $\mathcal{G}$ up to a phase; it maps the code space to itself and acts on it nontrivially. The *weight* of a Pauli string is the number of its factors different from I, and the *distance* of the code is the least weight of a logical error. A *gate* is a unitary on one or two qubits, tensored with the identity on the others, and a *circuit* is a finite product of gates. A *Clifford unitary* is a unitary U that maps Pauli strings to Pauli strings under conjugation, up to a phase; on labels, conjugation by U is a linear map that preserves ω.

The structure (V,ω) is the Fraïssé limit of the finite spaces with an alternating form over $\mathbb{F}_2$, that is, the unique countable homogeneous structure whose finite substructures are, up to isomorphism, exactly these spaces (see Variatio 4); the proof, and the model theory of (V,ω), are in Variatio 24.

Theorem 22.1. *The theories of the Rado graph and of (V,ω) are simple and unstable. In (V,ω), nonforking independence is linear independence over the base: for small subsets A, B and C of a monster model of the theory of (V,ω), the relation $A \mathop{\smile\!\!\!\!\!\!|}_C B$ of nonforking independence (see Variatio 6) holds if and only if* $\operatorname{span}(A \cup C) \cap \operatorname{span}(B \cup C) = \operatorname{span}(C)$.

We prove Theorem 22.1 in section P.3; for (V, ω) the idea is a gluing of linear functionals. By quantifier elimination (see Variatio 24), the type of an element over a small set D records the linear relations of the element over $\mathrm{span}(D)$ and the values of ω between the element and $\mathrm{span}(D)$, which form a linear functional on $\mathrm{span}(D)$; every such functional is realized in a monster model by an element outside $\mathrm{span}(D)$, since every finite alternating extension of a finite subspace of V embeds in V over the subspace. In the independence theorem (see Variatio 20), a model M and sets B and D with $B \mathop{\smile\!\!\!\!\!\!|}_M D$ are given, together with realizations c_1 and c_2 of one type over M such that $c_1 \mathop{\smile\!\!\!\!\!\!|}_M B$ and $c_2 \mathop{\smile\!\!\!\!\!\!|}_M D$. Writing MB for $M \cup B$, and similarly for other unions, we see that the functionals defined by c_1 on $\mathrm{span}(MB)$ and by c_2 on $\mathrm{span}(MD)$ agree on M, and that the intersection $\mathrm{span}(MB) \cap \mathrm{span}(MD)$ equals M, because $B \mathop{\smile\!\!\!\!\!\!|}_M D$; the two functionals therefore glue to one functional on $\mathrm{span}(MBD)$. If the common type is algebraic over M, its realization lies in M and is itself the amalgam; otherwise an element linearly independent over $\mathrm{span}(MBD)$ that realizes the glued functional is the amalgam, by quantifier elimination. Tuples are treated in the same way, coordinate by coordinate. The independence property comes from nondegeneracy: for linearly independent $a_1, \ldots, a_n$, every pattern of values $\omega(a_i, b)$ is realized by some b.

For the physicist or operator theorist. In operator terms, the type of a Pauli string over a set of strings records which of them it commutes with, and, by nondegeneracy of ω, every pattern of commutation with strings whose labels are linearly independent is attained; this freedom is the independence property, and the independence theorem is the gluing of two such patterns that agree on a common subspace.

For the logician. On a qubit, the bit flip X exchanges the standard basis vectors $|0\rangle$ and $|1\rangle$ of $\mathbb{C}^2$, the phase flip Z multiplies $|1\rangle$ by -1, and the matrix $Y = iXZ$ does both, up to a phase; a Pauli string on n qubits applies one of these operations, or the identity, to each qubit. Pauli strings serve both as measurements, with outcomes ± 1, and as errors, the operations that noise applies to a register; every operator on n qubits is a linear combination of them, which is why the restriction to Pauli errors loses nothing. As an error, a Pauli string P acts on density operators by $\rho \mapsto P\rho P^*$, an action that does not change when P is multiplied by a phase and is therefore determined by the label of P in $\mathbb{F}_2^{2n}$; the form ω records which pairs of strings can be measured jointly: those whose labels v and w satisfy $\omega(v, w) = 0$.

The Rado graph also lives inside the Pauli group, as a pattern of commutation between bit flips and idealized Pauli strings (see Variations 27 and 28). Like the Rado graph, the structure (V, ω) is simple and unstable and has the independence property; its nonforking independence, however, is linear independence over the base, and its independence theorem becomes the gluing of linear functionals.

VARIATIO 23

a 2 Clav.

The dictionary of simplicity

Three notions of model theory (Fraïssé limits, forking and the independence theorem) are at work in quantum error correction under other names: in this variation we set out the third dictionary of the memoir (see Table 7), after the dictionaries of stability and of NIP (see Variations 3 and 13), and present the running example; the physical side of this dictionary was introduced in Variatio 22. In each of the next three variations we state a theorem, prove it, and read it twice. A single example runs through all four: the Laflamme–Miquel–Paz–Zurek five-qubit code [**LMPZ96**], the smallest code that corrects an arbitrary error on one qubit, which Bennett, DiVincenzo, Smolin and Wootters found independently [**BDSW96**].

For the two code-specific forking rows, fix a stabilizer code on finitely many qubits and a nonzero error label on these qubits with zero syndrome; a set of logical labels is *full* if, together with the stabilizer labels, it spans all zero-syndrome labels on the same qubits.

Example 23.1. The five-qubit code has the stabilizer generators

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

and stores 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, the string of four bits that records which of the generators $g_1, \ldots, g_4$ anticommute with it, and there are exactly fifteen nonzero syndromes of four bits: the code is *perfect*. For instance, the error $XIIII$ has syndrome $(0,0,0,1)$, the error $IIZII$ has syndrome $(0,0,1,0)$, and the error $YIIII$ has syndrome $(1,0,1,1)$.

Model theory	**Quantum error correction**	
the structure (V, ω)	finite-support Pauli strings modulo phases, with label addition and the alternating commutation form	Variatio 22
the age of (V, ω), its finite substructures up to isomorphism	finite alternating label spaces, including their linear relations	Proposition 24.1
ultrahomogeneity	Clifford circuits; encoding circuits of codes	Theorem 24.2
algebraic closure of a set of labels	its linear span: products of the corresponding Pauli strings modulo phases	Variatio 22
the type of an error label over a set of labels	its syndrome and all its linear relations over the set	Theorem 25.1
the type over known labels forks over base labels among them	the error is a sum of known labels but not a sum of base labels	Theorem 25.1
the type over stabilizer labels does not fork over $\emptyset$	the fixed zero-syndrome error is logical	Corollary 25.2
the type over stabilizer and logical labels forks over the stabilizer labels	the same error is logical: the full set of logical labels makes it a sum of known labels	Theorem 25.1
the independence theorem	when the check spans meet only in the common base span, syndrome functionals agreeing there have a joint extension	Proposition 26.1

TABLE 7. The dictionary of simplicity: the Pauli group as a structure.

VARIATIO 24

a 1 Clav. Canone all'Ottava

Fraïssé limits and Clifford circuits

Two properties make the labels of Pauli strings, with their commutation form, a Fraïssé limit: every finite pattern of commutation occurs among them, and every isomorphism between two of their finite substructures extends to an automorphism. In Proposition 24.1 below we give the first a physical form, showing that, for every symmetric matrix of zeros and ones with zeros on the diagonal, there are Pauli strings, tensor products of Pauli matrices, one for each row of the matrix, such that two of them anticommute exactly when the matrix has a one in the corresponding entry; in Theorem 24.2 we give the second, showing that any two families of Pauli strings with linearly independent labels and the same pattern of commutation are carried to each other by a Clifford circuit, a product of gates that map Pauli strings to Pauli strings under conjugation, up to a phase. The two readings coincide: for the logician, the Pauli group is a Fraïssé limit; for the physicist, every isometry between its finite parts is implemented by a Clifford circuit.

In the language with $+$, 0 and a relation for $\omega(x, y) = 1$, where ω is the commutation form of Variatio 22, the finite spaces with an alternating form amalgamate freely, every finite one embeds in a nondegenerate one, and Witt's theorem extends every isometry between subspaces of a finite nondegenerate space, that is, every linear bijection between them that preserves the form, to an isometry of the whole space. A back-and-forth argument then shows that the structure (V, ω) of labels of Pauli strings on countably many qubits is the Fraïssé limit of the finite spaces with an alternating form over $\mathbb{F}_2$, the unique countable homogeneous structure whose finite substructures are, up to isomorphism, exactly these spaces (see Variatio 4).

A theory is $\aleph_0$-*categorical* if all its countably infinite models are isomorphic; by the Engeler–Ryll-Nardzewski–Svenonius theorem, for a complete theory in a countable language with infinite models this holds exactly when, for each positive integer n, the n-tuples of every model, for instance of the ultrapower $V^{\mathcal{U}}$ along a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, realize only finitely many types. The theory of (V, ω) is $\aleph_0$-categorical and eliminates quantifiers (see Variatio 1): two tuples, in V or in $V^{\mathcal{U}}$, have the same type as soon as they satisfy the same linear relations and ω takes the same values on corresponding pairs of their entries. Algebraic closure in (V, ω) is linear span. An $\aleph_0$-categorical structure is *smoothly approximable* if it is the union of a chain $F_1 \subseteq F_2 \subseteq \cdots$ of finite substructures such that, for every n, any two tuples of F_n with the same type are conjugate by an automorphism that maps F_n onto itself; for (V, ω), the sets of labels of the strings on the first n qubits form such a chain, and Cherlin and Hrushovski count (V, ω) among the basic smoothly approximable structures [**CH03**]. The next two results make these facts physical.

Proposition 24.1. *Let m be a positive integer. If Θ is a symmetric $m \times m$ matrix over $\mathbb{F}_2$ with zero diagonal, then there are Pauli strings $P_1, \ldots, P_m$ with linearly independent labels, on at most m qubits, such that P_i and P_j anticommute exactly when $\Theta_{ij} = 1$.*

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

Theorem 24.2. *Let n be a positive integer, let W and W' be subspaces of the space $\mathbb{F}_2^{2n}$ of labels of Pauli strings on n qubits, and let $f\colon W \to W'$ 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 such that, for every $v \in W$, the Pauli string P_v with label v satisfies $UP_vU^* = \pm P_{f(v)}$. If, moreover, the vectors $v_1, \ldots, v_m$ 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. The idea is to extend f to an isometry of the whole space, to implement the extension by a Clifford unitary, and to correct the signs by a Pauli string. By Witt's theorem [**Art57**], the map f extends to a linear bijection T of $\mathbb{F}_2^{2n}$ that preserves ω. Every such T is induced by conjugation by a Clifford unitary [**CRSS98**, **DDM03**]; let U be one. Since UP_vU^* 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_bU has the required property. □

For integers n and k with $0 \leq k \leq n$, the *trivial code* that stores k qubits in n is the code whose stabilizer group is generated by the $n-k$ strings that act as Z on one of the first $n-k$ qubits and as the identity on the others.

Corollary 24.3. *Let n be a positive integer. Any two stabilizer groups on n qubits with the same number of independent generators are conjugate by a Clifford unitary. In particular, for every integer k with $0 \leq k \leq n$, every stabilizer code that stores k qubits in n has an encoding circuit made of Clifford gates, which maps the code space of the trivial code onto the code space of the given code.*

PROOF. Let $g_1, \ldots, g_m$ and $g'_1, \ldots, g'_m$ be independent generators of the two groups, with labels v_i and v'_i, respectively. Both families span isotropic subspaces, so the linear map $v_i \mapsto v'_i$ preserves ω, and Theorem 24.2 yields a Clifford unitary U with $UP_{v_i}U^* = \pm P_{v'_i}$. The signs of the generators g_i and g'_i are then matched by a Pauli correction, as in the proof of the theorem. □

On countably many qubits, every finite-dimensional subspace of V lies in the labels of the first n qubits for some n. Applying Theorem 24.2 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 or operator theorist. The homogeneity of (V, ω) says that the automorphisms of the structure, on any finite set of labels, are implemented by unitaries: every isometry between finite commutation patterns of Pauli strings is induced by a Clifford circuit, and the encoding circuit of a stabilizer code is one such isometry (see Corollary 24.3).

For the logician. The Clifford unitaries are, up to phases, the products of three kinds of gates: the Hadamard gate $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right)$, which exchanges the Pauli matrices X and Z under conjugation; the phase gate $\mathrm{diag}(1, i)$, which sends X to Y and fixes Z; and the controlled-NOT gate on two qubits, which flips the second qubit when the first is in the basis state $|1\rangle$. Conjugation by a Clifford unitary maps Pauli measurements to Pauli measurements and Pauli errors to Pauli errors, which is why these gates are the basic operations of stabilizer error correction. An *encoding circuit* for a code that stores k qubits in n is a unitary that maps the code space of the trivial code, in which the first $n - k$ qubits are fixed in $|0\rangle$ and the last k carry the information, onto the code space of the code. By Corollary 24.3, every stabilizer code has one, built from these gates.

VARIATIO 25

a 2 Clav. Adagio

Forking and logical errors

Some errors of a stabilizer code commute with every check, that is, with each of the generators of the stabilizer group that the code measures; every check then returns the outcome 1, and the syndrome, the list of these outcomes read as bits, is zero. Such errors are undetectable, and the decoder, the procedure that infers a correction from the syndrome, does not correct them. Some undetectable errors are harmless, others corrupt the stored qubits, and the syndrome does not distinguish the two kinds.

In this variation we separate them by forking (see Variatio 6): for an undetectable error other than the identity, the type of its label over the labels of the stabilizer group forks over the empty set exactly when the error is harmless (see Corollary 25.2). We compute forking for the Pauli group from Theorem 22.1, read forking in a code, and count the two kinds of error in the five-qubit code.

Theorem 25.1. *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 $\operatorname{tp}(e/B)$ forks over D if and only if $e \in \operatorname{span}(B) \setminus \operatorname{span}(D)$.*

Proof. By Theorem 22.1, the type $\operatorname{tp}(e/B)$ does not fork over D exactly when $\operatorname{span}(eD) \cap \operatorname{span}(B) = \operatorname{span}(D)$. If $e \in \operatorname{span}(D)$, then $\operatorname{span}(eD) = \operatorname{span}(D)$, and the equality holds. If $e \notin \operatorname{span}(D)$, then we have $\operatorname{span}(eD) = \operatorname{span}(D) \oplus \{0, e\}$, and the equality fails exactly when some vector $e + d$ with $d \in \operatorname{span}(D)$ lies in $\operatorname{span}(B)$, that is, exactly when $e \in \operatorname{span}(B)$. □

Forking is thus a strong condition: the type of e over B forks over D exactly when e is a linear combination of elements of B, that is, up to a phase, a product of the corresponding Pauli strings, without being a linear combination of elements of D. Prescribed commutation values never cause forking: for $D \subseteq B$, a type over D of an element outside $\operatorname{span}(D)$ extends, by any prescription of commutation values on B consistent with linearity and with the type over D, to a type over B that does not fork over D, provided the realization is chosen outside $\operatorname{span}(B)$. In particular, the type over actual labels of an element that an ultrapower adds to the space of labels never forks over the empty set. Indeed, let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$, and let $(e_n)_{n \in \mathbb{N}}$ be a sequence of labels in the space V of labels of Pauli strings on countably many qubits (see Variatio 22) whose class $[e_n]_{\mathcal{U}}$ in the ultrapower $V^{\mathcal{U}}$ lies outside V, as it does when the labels e_n are pairwise distinct; then the class does not lie in the span of any subset B of V, and by Theorem 25.1 the type of the class over B does not fork over $\emptyset$. In a stabilizer code this criterion separates the two kinds of undetectable error other than the identity.

Corollary 25.2. *Let n be a positive integer, and regard $\mathbb{F}_2^{2n}$ as the subspace of V formed by the labels of the strings on the first n qubits. Let $\mathcal{G}$ be a stabilizer group on n qubits with set*

of labels S, and let $e \in \mathbb{F}_2^{2n}$ be a nonzero label with zero syndrome. Then the type $\mathrm{tp}(e/S)$ *forks over $\emptyset$ if and only if $e \in S$.*

PROOF. Apply Theorem 25.1 with $D = \emptyset$ and $B = S$; the span of S is S itself, and $\mathrm{span}(\emptyset) = \{0\}$. □

Thus the type over S of a nonzero label with zero syndrome forks over $\emptyset$ exactly when the label is harmless (see Figure 11). Since a logical error is a nonzero label with zero syndrome outside S, the distance of a code that encodes at least one qubit is the least weight of a nonzero label in $\mathbb{F}_2^{2n}$ with zero syndrome whose type over S does not fork over $\emptyset$.

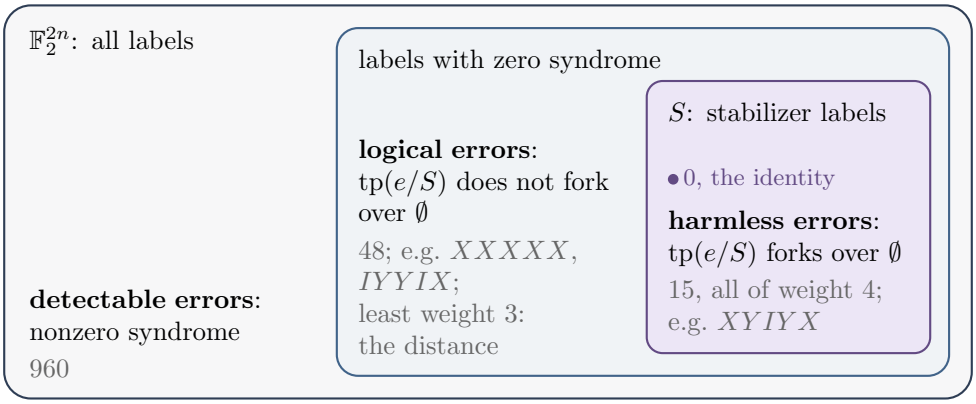


FIGURE 11. The labels of a stabilizer code, modulo phases and nested by inclusion, with the counts for the five-qubit code.

In the five-qubit code (see Example 23.1), consider two errors, up to phases: the product $g_1 g_2 = XYIYX$ of the first two stabilizer generators and the logical operator $\bar{X} = XXXXX$. Both have zero syndrome. The first lies in the stabilizer group, so its type over S forks over $\emptyset$, and it is harmless. The second does not, so its type does not fork over $\emptyset$, and it flips the stored qubit. The least weight of an error of the second kind is 3, as the count below shows, attained for instance by $\bar{X} g_1 = IYYIX$ up to a phase: the code has distance 3.

We now count the undetectable errors of the five-qubit code. For the $[[5, 1, 3]]$ code on five qubits there are $4^5 = 1024$ Pauli strings up to phase. The undetectable ones, those with zero syndrome, are the labels on which ω vanishes against the four checks; they form a subspace of dimension $10 - 4 = 6$, sixty-four labels in all. Sixteen of them lie in the stabilizer group. One is the identity, which is no error at all, and whose type over S does not fork over $\emptyset$, since it lies in $\mathrm{span}(\emptyset)$. The other fifteen are strings of weight four; their types over S fork over $\emptyset$, and they are harmless. The remaining forty-eight are logical errors, and their types do not fork over $\emptyset$: thirty of weight three, eighteen of weight five, and none of weight four.

weight of the nonidentity undetectable error	3	4	5
harmless: the type over S forks over $\emptyset$	0	15	0
logical: the type over S does not fork over $\emptyset$	30	0	18

Every undetectable error of weight four is harmless. Every undetectable error of weight three is logical. The distance of the code, three, is the weight at which the nonforking errors begin.

Forking also tracks what is learned as labels are added to the parameters. Suppose that the label of the logical operator $\bar{Z} = ZZZZZ$ of the five-qubit code is added. The type of $\bar{X}$ over $S \cup \{\bar{Z}\}$ records that $\bar{X}$ anticommutes with $\bar{Z}$, so the enlarged set of parameters separates $\bar{X}$ from the logical errors that commute with $\bar{Z}$. The addition is a change of parameters, not a new check of the code: a measurement of $\bar{Z}$ reads the stored qubit, and on the logical state $|+\rangle$, which $\bar{X}$ fixes, the outcomes of such a measurement have the same distribution with or without the error. The type of $\bar{X}$ over $S \cup \{\bar{Z}\}$ still does not fork over S, because $\bar{X}$ is not a product of $\bar{Z}$ and the stabilizers. By Theorem 25.1, the type of $\bar{X}$ over a set $B \supseteq S$ forks over S exactly when $\bar{X}$ lies in $\mathrm{span}(B)$. As in Variatio 23, a set of logical operators is full if its labels, together with the stabilizer labels, span all zero-syndrome labels on the code's qubits. After all these labels have been added, the type of every logical error over the enlarged parameter set forks over the stabilizer labels, by Theorem 25.1: the error lies in the enlarged span but outside the stabilizer span.

Forking can be observed directly in an elementary extension. Take $D = \emptyset$, $B = S$, and the harmless error $e = g_1 g_2$, whose label is $g_1 + g_2$. The formula $x = g_1 + g_2$, with parameters g_1, g_2, divides over $\emptyset$: on fresh qubits of the structure on countably many qubits there are pairs $(g_1^{(j)}, g_2^{(j)})$, $j \in \mathbb{N}$, with the same commutation pattern as (g_1, g_2) and pairwise disjoint supports, hence with pairwise distinct sums; by Ramsey's theorem and compactness they may be taken indiscernible over $\emptyset$, and no label equals two of the sums. The type $\mathrm{tp}(e/S)$ therefore forks over $\emptyset$, as Corollary 25.2 states. For a logical error e, by contrast, no formula of $\mathrm{tp}(e/S)$ divides over $\emptyset$, in accordance with Theorem 25.1: by quantifier elimination, the type says only that $e \notin \mathrm{span}(S)$ and that $\omega(s, e) = 0$ for $s \in S$, and it is realized by labels on qubits disjoint from any given finite set.

In the space V of labels on countably many qubits (see Variatio 22), the type of a logical error over S has several nonforking extensions: all labels outside $\mathrm{span}(S)$ with zero syndrome have the same type over S, since, by quantifier elimination, the type over S of such a label e is determined by the absence of linear relations between e and $\mathrm{span}(S)$ and by the linear functional $\omega(\cdot, e)$ on $\mathrm{span}(S)$, which is 0. Its nonforking extensions to $S \cup \{\bar{Z}\}$ are at least two: the extension realized by $\bar{X}$, which anticommutes with $\bar{Z}$, and the extension realized by a label supported on a sixth qubit, which commutes with $\bar{Z}$; neither label lies in $\mathrm{span}(S \cup \{\bar{Z}\})$, so neither extension forks over S. The second realization leaves the five qubits of the code: on these qubits the zero-syndrome labels form the direct sum of S and the span of the labels of $\bar{X}$ and $\bar{Z}$, so a zero-syndrome label that commutes with $\bar{Z}$ lies in $\mathrm{span}(S \cup \{\bar{Z}\})$, and its type over $S \cup \{\bar{Z}\}$ is algebraic. The several nonforking extensions are therefore a feature of the ambient structure (V, ω). There, learning the values of ω at further labels gives only conditions on commutation, which never cause forking, and no finite set of such values determines a label; a decoder works with more, namely the qubits that errors can affect and their probabilities, which the type does not record. Forking occurs only when the parameters come to contain the error as a product: if $e = b_1 + \cdots + b_m$ with $b_i \in B$ and $e \notin \mathrm{span}(D)$, then the formula $x = b_1 + \cdots + b_m$ divides over D, by the argument given above for $g_1 g_2$, and with $D = S$ the type over B then determines the error.

For the physicist or operator theorist. In operator terms, for a set of known Pauli strings that contains a set of base strings, the type of an error over the known strings forks over the base strings exactly when the error is, up to a phase, a product of known strings but not of base strings; in the ambient label structure, knowledge of which known strings commute with the error never causes forking.

For the logician. An error on a code with stabilizer group $\mathcal{G}$ is undetectable when it commutes with every element of $\mathcal{G}$, and then it maps the code space to itself. If it lies in $\mathcal{G}$ up to a phase, it acts on the code space as a scalar and the stored information is intact; otherwise it is a logical error: it acts as a nontrivial unitary on the stored qubits, and no measurement of elements of $\mathcal{G}$ tells the two cases apart, since both yield the trivial syndrome. In Corollary 25.2 the distinction is read as forking: for an undetectable error other than the identity, the type of its label over the labels of $\mathcal{G}$ forks over $\emptyset$ exactly when the error is harmless. In the five-qubit code, every nonidentity error on at most two qubits has a nonzero syndrome, every error on one qubit has its own syndrome and can therefore be corrected (see Example 23.1), and the distance, 3, is the least weight of a logical error.

Forking thus draws exactly the distinction that the syndrome cannot: an undetectable error other than the identity is harmless if and only if the type of its label over the labels of the stabilizer group forks over the empty set. In the next variation we turn to the independence theorem, the other half of the Kim–Pillay characterization of simplicity (see Variatio 20).

VARIATIO 26

a 2 Clav.

The gluing of syndromes

In this variation we show that the independence theorem, read in the Pauli group, says that a single error produces two syndromes prescribed for two sets of checks designed independently, provided the syndromes agree on the common checks (see Proposition 26.1 below). Two sets of checks, Pauli strings measured to detect errors, on the same finitely many qubits are *designed independently* when the spans of their labels meet only in the span of the labels of their common checks. The independence theorem (see Variatio 20) amalgamates nonforking extensions of one type over a model to two sets independent over the model; in the Pauli group its counterpart holds over the span of the labels of the common checks, an algebraically closed set, in place of a model.

Proposition 26.1. *Let n be a positive integer, let A_0, B and D be subsets of the space $\mathbb{F}_2^{2n}$ of labels of Pauli strings on n qubits, suppose that $\operatorname{span}(B \cup A_0) \cap \operatorname{span}(D \cup A_0) = \operatorname{span}(A_0)$, and let f_1 and f_2 be linear functionals on $\operatorname{span}(B \cup A_0)$ and on $\operatorname{span}(D \cup A_0)$ that agree on $\operatorname{span}(A_0)$. Then some label $e \in \mathbb{F}_2^{2n}$ has the commutation values $\omega(s, e) = f_1(s)$ for $s \in \operatorname{span}(B \cup A_0)$ and $\omega(s, e) = f_2(s)$ for $s \in \operatorname{span}(D \cup A_0)$.*

Proof. Since f_1 and f_2 agree on the intersection of their domains, there is a linear functional f on $\operatorname{span}(B \cup A_0) + \operatorname{span}(D \cup A_0)$ extending both. This functional extends to $\mathbb{F}_2^{2n}$, and, because ω is nondegenerate, every linear functional on $\mathbb{F}_2^{2n}$ has the form $\omega(\cdot, e)$ for a unique e. □

In the five-qubit code, with the stabilizer generators $g_1, \dots, g_4$ of Example 23.1, let $A_0 = \emptyset$, $B = \{g_1, g_2\}$ 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 23.1, even by an error on at most one qubit, such as $YIIII$ for the syndrome $(1, 0 \,|\, 1, 1)$. If D is replaced by $\{g_1 g_2\}$, the hypothesis fails, and so does the conclusion: the syndrome bit of $g_1 g_2$ is the sum of those of g_1 and g_2. The proposition glues linear commutation data; for the two families of checks to be measured together on one code, the checks of their union must also commute with one another, which the hypothesis on spans does not ensure.

On a finite number n of qubits, the information carried by a syndrome can be measured by counting. If $s_1, \dots, s_m$ are linearly independent labels in $\mathbb{F}_2^{2n}$, then, for every syndrome $\mathbf{b} \in \mathbb{F}_2^m$, exactly 2^{2n-m} labels e satisfy $\omega(s_i, e) = \mathbf{b}_i$ for all $i \le m$, since the map $e \mapsto (\omega(s_i, e))_{i \le m}$ is linear and onto. A uniformly random error has a uniformly random syndrome, which reveals exactly m bits. The count records codimension, which forking ignores, and it reflects the smooth approximability of (V, ω), the structure of labels on countably many qubits, by finite structures (see Variatio 24).

For the physicist or operator theorist. In operator terms, consider two sets of Pauli strings whose labels span subspaces meeting only in the span of their common part. By

Proposition 26.1, two prescriptions of commutation values on them are realized by a single Pauli string when each prescription is consistent with the linear relations among its strings and the two agree on the common part.

For the logician. A syndrome is the list of outcomes, each ± 1, of measuring the checks of a code on the corrupted state, and it records for each check whether the error anticommutes with the check. By Proposition 26.1, if two families of checks are independent over their common part, in the sense of the hypothesis on spans, then any two lists of outcomes, each produced by some Pauli error and agreeing on the common checks, are produced by a single Pauli error; consequently, once the outcomes on the common checks are fixed, no outcome of one family excludes an outcome of the other. In the five-qubit code, the outcomes of g_1, g_2 and of g_3, g_4 can be prescribed separately, whereas for the checks g_1, g_2 and $g_1 g_2$ the third outcome is the product of the first two.

The independence theorem is thus to simplicity what uniqueness of the independent product (see Theorem 6.4) is to stability: the property of amalgams that characterizes the class.

VARIATIO 27

a 2 Clav. Canone alla Nona

The Rado graph inside the Pauli group

Every graph on a finite set of qubits determines a quantum state, its graph state. In Proposition 27.1 below, we show by means of graph states that every relation between two finite sets is a pattern of commutation between Pauli strings, and in Variatio 28 that, with idealized strings on one side, so is every relation between countable sets, the Rado graph among them. The Pauli group carries, in addition, the linear structure of its labels, which mirrors the multiplication of Pauli strings, and in Variatio 29 we show that this structure lets actual preparations witness the loss of simplicity under mixing.

A *stabilizer state* on n qubits is a unit vector fixed by every element of a stabilizer group (see Variatio 22) with n independent generators; the vectors fixed by such a group form a one-dimensional space, the code space of the group. Let G be a graph on a finite set W of qubits. The *graph state* of G is the stabilizer state, unique up to a phase, of the stabilizer group with the generators

$$K_v = X_v \prod_{w \sim v} Z_w \qquad (v \in W),$$

where the strings X_v and Z_w act as X on the qubit v and as Z on the qubit w, respectively, and as the identity on the other qubits, and where the product runs over the neighbors w of v in G.

Proposition 27.1.

(a) *For every finite graph G on W and all $u, v \in W$, the strings X_u and K_v anticommute exactly when u and v are adjacent.*

(b) *For every relation $\varphi \subseteq A \times B$ between finite sets, there are Pauli strings P_a, $a \in A$, and Q_b, $b \in B$, such that P_a and Q_b anticommute exactly when $(a, b) \in \varphi$.*

Proof. (a) The string X_u commutes with every X factor of K_v, and with every Z factor on a qubit other than u; it anticommutes with Z_u, which is a factor of K_v exactly when $u \sim v$. (b) We place the elements of A on distinct qubits, let $P_a = X_a$, and let Q_b be the product of Z_a over the a with $(a, b) \in \varphi$. □

For a finite graph, the strings K_v of part (a) are actual Pauli strings, the generators of its graph state.

The model-theoretic reading. Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$. In the ultrapower $V^{\mathcal{U}}$ of the space V of labels of Pauli strings on countably many qubits, the class $[b_n]_{\mathcal{U}}$ of a sequence $(b_n)_{n \in \mathbb{N}}$ of labels defines on V the linear functional $v \mapsto \lim_{n,\mathcal{U}} \omega(v, b_n)$, where ω is the commutation form; this functional records the ω-type of the class over V, and every linear functional f on V arises in this way, since, by nondegeneracy, the restriction of f to the labels of the strings on the first n qubits has the form $\omega(\cdot, b_n)$ for some label b_n of a string on these qubits. Applied to the labels of the bit flips X_a, for a in a countable set A of qubits, this shows that every relation between A and a countable set B is realized as

the pattern of commutation between the flips X_a and suitable idealized labels, one for each element of B, which is how the Rado graph lives inside the Pauli group. The two theories nevertheless differ: algebraic closure is trivial in the Rado graph and is linear span in (V, ω), and accordingly the type of a vertex forks only when the vertex is named (see Variatio 21), while the type of a label over a set of labels forks over a subset exactly when the label is a linear combination of labels of the set without being a linear combination of labels of the subset (see Theorem 25.1).

For the physicist or operator theorist. In operator terms, every relation between two finite sets is a pattern of anticommutation between bit flips and products of phase flips, and, for a finite graph, these products, multiplied by bit flips, are the stabilizers of its graph state.

For the logician. A graph state is the unique state, up to a phase, left unchanged by every K_v; it is prepared by placing every qubit in $|+\rangle$ and applying, for each edge $\{v, w\}$, the controlled-Z gate on the qubits v and w, which multiplies by -1 the basis vector $|1\rangle_v|1\rangle_w$, in which both qubits are in the state $|1\rangle$, and fixes the other three basis vectors of the pair. Measuring the stabilizers K_v after a bit flip X_u gives the syndrome of the flip: by Proposition 27.1(a), the outcome is -1 exactly for the neighbors v of u, so the syndrome of a single flip is a row of the adjacency matrix, and the flip is detected exactly when u has a neighbor: the generator of an isolated vertex u is X_u itself, which fixes the graph state. For the Rado graph on countably many qubits, each K_v would act on infinitely many qubits, which is why Bob's strings become idealized in Variatio 28, while the flips X_u remain actual.

The Pauli group thus carries, beyond the commutation patterns that it shares with the Rado graph, the linear structure of its labels; in Variatio 29 we show that this structure decides whether the loss of simplicity under mixing is already visible among actual preparations.

VARIATIO 28

a 2 Clav.

Mixing destroys simplicity

In this variation we show that mixing destroys simplicity: for diagonal observables, mixtures turn the independence property into the tree property of the second kind (see Theorem 28.1), so that the simple theories of the Pauli and Rado label structures give rise to nonsimple theories of their mixture-and-probe structures. The structures, and hence the theories being compared, have changed. The collapse is the third effect of mixing recorded in Table 1, after one in each of the first two parts: in Part I, for stable observables, finite mixtures of actual preparations simulate idealized ones, and in Part II mixing makes hereditary separability of the idealized responses imply their metrizability. The idea of the proof is a comparison of two averages: the difference between the responses to the uniform mixtures on two complementary halves of a block of the labels that Alice's register can hold measures how a probe weighs the two halves, and halves meeting in a quarter of the block cannot both give a response difference greater than one half.

Before simplicity can be tested on mixtures, the mixtures must become elements of a structure; the diagonal observable of a relation provides one. Let $\varphi\colon A \times B \to \{0,1\}$ be a relation between countable sets of labels, and consider its diagonal observable (see Variatio 14)

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

Alice's register holds a label a, Bob's register holds a label b, and the observable reads $\varphi(a,b)$. A mixture of Alice's preparations is given by a probability measure μ on A. The probes form the space $\mathcal{P}(K_\varphi)$ of Radon probability measures on the closure K_φ of the functions $\varphi(\cdot,b)$ in $\{0,1\}^A$, the idealized labels of Bob (see Lemma P.63). The proof of the lemma identifies the probe algebra $\mathfrak{A}_{C_\varphi}$ of Variatio 11 with $C(K_\varphi)$, by the Stone–Weierstrass theorem, and the probe space Y_{C_φ} with $\mathcal{P}(K_\varphi)$, by the Riesz representation theorem. The response of the mixture μ to a probe $\nu \in \mathcal{P}(K_\varphi)$ is the probability $R(\mu,\nu) = \int_{K_\varphi} \sum_{a \in A} \mu(\{a\})\lambda(a)\, d\nu(\lambda)$ that the observable reads 1, and we write $\mathfrak{r}_\mu\colon \mathcal{P}(K_\varphi) \to [0,1]$ for the response function defined by $\mathfrak{r}_\mu(\nu) = R(\mu,\nu)$. The observable thus defines a two-sorted structure $\mathcal{M}_\varphi$ of real-valued logic, with the mixtures of Alice and the probes of Bob as sorts and the response R as predicate, and we call the observable *simple* if the theory of $\mathcal{M}_\varphi$, a real-valued structure in the sense of Keisler [**Kei23**], is simple in the sense of Ben Yaacov [**BY03**]: no formula $\theta(x,y)$ has the tree property, that is, for no real $\varepsilon > 0$ and integer $k \geq 2$ is there a tree of parameters such that the conditions $\theta(x,d) = 0$ along each branch are jointly satisfiable while, for any k siblings $d_1, \dots, d_k$, the conditions $|\theta(x,d_1)| \leq \varepsilon$, …, $|\theta(x,d_k)| \leq \varepsilon$ are not. A formula is *affine* if it is built from atomic formulas with the quantifiers and the connectives $u(t_1,\dots,t_n) = c_0 + \sum_{i=1}^n c_i t_i$, where $n \geq 1$ and $c_0, \dots, c_n$ are real numbers. In an elementary extension, for instance in the ultrapower $\mathcal{M}_\varphi^{\mathcal{U}}$ along a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the new elements of Alice's sort are idealized preparations of Alice.

The *commutation observable* of the Pauli group is the diagonal observable of the commutation form ω on labels, and the *Rado observable* is the diagonal observable of the edge relation E of the Rado graph. Both have the independence property already on pure labels. On countably many qubits, Proposition 27.1(b) extends to countable sets A and B with the strings Q_b replaced by idealized ones: with the elements of A placed on distinct qubits, as in its proof, for each $b \in B$ some linear functional on the space V of labels takes the value 1 at the label of the bit flip X_a exactly when $(a, b) \in \varphi$, and the compact group V^* of all linear functionals on V is the space of idealized labels of Bob for the commutation observable (see Lemma P.65). In particular, the Rado graph and the random bipartite graph, the Fraïssé limit of the finite bipartite graphs with the two sides as two sorts, are realized as patterns of commutation between single-qubit flips and idealized Pauli strings, while for a finite graph Bob's strings are actual Pauli strings, the generators of its graph state.

The pure and mixed responses have the independence property together (see Lemma P.68). Here both properties concern Alice's responses, first on K_φ, whose points, identified with the point masses at them, we call *pure probes*, and then on the probe space $\mathcal{P}(K_\varphi)$. The nontrivial direction, from NIP for the pure responses to NIP for the mixed ones, rests on Rosenthal's dichotomy and Stegall's theorem on absolutely convex hulls. Conversely, pure preparations and pure probes are among the mixed preparations and the probes; finite zero/one patterns on K_φ meet the dense actual labels, so the hypothesis of Theorem 28.1 below is equivalent to failure of NIP for the observable.

Theorem 28.1. *Let C_φ be a diagonal observable, and let $a_0, a_1, \dots$ be distinct labels such that every finite pattern of values $\varphi(a_n, b)$ is realized by some label b. Then the affine formula $1 - R(y_1, x) + R(y_2, x)$ has the tree property of the second kind in the theory of $\mathcal{M}_\varphi$. In particular, a simple diagonal observable has no independence property.*

Proof. We first build finite arrays of every size, whose entries are pairs of uniform mixtures on complementary halves of blocks of labels, and then pass to an infinite array in an ultrapower. Let a positive integer N be given, and choose an integer t with $2^t > N$. We divide $N2^t$ of the labels a_n into N blocks $S_0, \dots, S_{N-1}$, identify each block with $\mathbb{F}_2^t$, choose distinct nonzero $w_0, \dots, w_{N-1} \in \mathbb{F}_2^t$, and, for $i, j < N$, let $T_{ij} = \{v \in S_i : w_j \cdot v = 1\}$, a half of the block S_i; two different halves in a block meet in a quarter of it. Let $\mu_{ij}^{\pm}$ be the uniform mixtures on T_{ij} and on $S_i \setminus T_{ij}$, respectively. For a probe ν with probabilities q_a of answering 1 at the label a, and with $m = 2^t$, we have

$$R(\mu_{ij}^+, \nu) - R(\mu_{ij}^-, \nu) = \frac{2}{m}\Big(\sum_{a \in T_{ij}} q_a - \sum_{a \in S_i \setminus T_{ij}} q_a\Big) \le 1,$$

and for $j \neq j'$ the sum of the two values in row i is $\frac{4}{m}\big(\sum_{T_{ij} \cap T_{ij'}} q_a - \sum_{S_i \setminus (T_{ij} \cup T_{ij'})} q_a\big) \le 1$, so the two values are never both greater than $1/2$. For any choice of one j in each row, the independence property provides a label b that answers 1 exactly on the chosen halves, and at b every chosen value is 1. These finite arrays exist for every N. For each fixed pair (i, j), use its entry in the array of size N when $N > \max(i, j)$ and fill the remaining finitely many coordinates arbitrarily. The resulting sequences define an infinite array in $\mathcal{M}_\varphi^{\mathcal{U}}$, and Łoś's theorem transfers the row bound and the satisfiability of every finite path. More explicitly, for a path $f\colon \mathbb{N} \to \mathbb{N}$ the countable conditions

$$R(\mu_{if(i)}^+, x) - R(\mu_{if(i)}^-, x) = 1 \qquad (i \in \mathbb{N})$$

are finitely satisfiable in the ultrapower. Its $\aleph_1$-saturation therefore realizes the entire path. For the final assertion, failure of NIP for the mixed responses gives the independent sequence of actual labels by the transfer argument above. The construction then gives TP_2, so the observable is not simple. □

The theorem is the counterpart, for independent preparations, of the theorem of Ben Yaacov that the randomization of a theory with the independence property has the tree property of the second kind [**BY13**]; the randomization is Keisler's real-valued theory of random elements $\xi\colon \Omega \to M$ of a model M of the theory, defined on an atomless probability space $(\Omega, \mathbb{P})$, with predicates $\mathbb{P}[\theta(\xi_1, \dots, \xi_n)]$ for the formulas θ of the theory. The theorem also mirrors Theorem 19.1, by which a separable convex Rosenthal compactum that is hereditarily separable is metrizable: in both cases, a class of the classical theory disappears because states can be mixed. The tree of the proof lives in an elementary extension, among idealized preparations. Whether actual preparations already carry it separates the commutation observable from the Rado observable, as we show in the next variation.

For the physicist or operator theorist. In operator terms, the theorem concerns uniform mixtures on halves of blocks of labels: no probe distinguishes two halves of one block, which meet in a quarter of it, from their complements with response differences both greater than one half, while every choice of one half in each block is distinguished from its complement, with difference one, by a single label of Bob.

For the logician. A diagonal observable reads classical data stored in two quantum registers: Alice's register holds a label a as the basis vector $|a\rangle$, Bob's holds b as $|b\rangle$, and the measurement returns $\varphi(a, b)$; on a density operator only the probabilities $\langle a, \rho a\rangle$ of the labels matter, so Alice's preparations are, in effect, random labels. For the commutation observable, a label of Alice is read as a check, a Pauli string that a code measures, and a label of Bob as an error, and the outcome of the measurement, 1 when the two anticommute and 0 otherwise, is the syndrome bit of the error at the check. The proof of Theorem 28.1 uses only uniform mixtures over halves of blocks of labels, which a laboratory prepares by choosing a label of the half uniformly at random, and the tree property of the second kind is a statement about the probabilities with which such random checks detect a fixed error.

In Part I mixing stood in for idealization; in Part II it supplied the midpoints of split pairs, such as the two one-sided idealizations of a particle at a point of the line in the comparison of positions; here it builds a tree.

VARIATIO 29

a 1 ô vero 2 Clav.

Actual and idealized trees

In this variation we show that the commutation observable and the Rado observable, which reach the same conclusion through the collapse theorem (Theorem 28.1), differ in where a witnessing array can be found. The commutation observable is the diagonal observable of the commutation form of the Pauli group, and the Rado observable is the diagonal observable of the edge relation of the Rado graph; for both, mixtures turn the independence property of the pure responses into the tree property of the second kind, in an elementary extension, among idealized preparations. For the commutation observable, all array parameters can be actual finite mixtures; for the Rado observable, no fixed quantifier-free formula in one probe variable has a witnessing array with actual preparation parameters.

In the theorem below, a sequence of instances is *uniformly k-inconsistent*, for an integer $k \geq 2$, if there are fixed reals $\alpha < \beta$ such that every instance takes a value at least β somewhere, but no probe gives values greater than α to any k distinct instances. For the commutation observable, finite parts of paths are realized by errors of finite support, while entire paths are realized by pure probes, points of the compact group V^* of idealized labels (see Variatio 28), which may correspond to Pauli strings of infinite support.

Theorem 29.1.

(a) *For the commutation observable, actual preparations witness the tree property of the second kind. Let $(g_{im})_{i,m\in\mathbb{N}}$ be linearly independent labels in V, and, for $i, j \in \mathbb{N}$, let μ_{ij} be the uniform mixture of Alice's preparations on the coset $g_{ij} + \operatorname{span}(g_{im} : m < j)$. Then, for every label $b \in V$ of Bob, the response $R(\mu_{ij}, b)$ equals 1 exactly when g_{ij} is the first of the labels $g_{i0}, g_{i1}, \ldots$ at which $\omega(\cdot, b)$ takes the value 1; for every choice of one entry in each of finitely many rows of the array $(\mu_{ij})_{i,j\in\mathbb{N}}$, there is a single label of Bob to which every chosen entry has response 1; and, for $j \neq j'$, no probe yields responses greater than $2/3$ to both μ_{ij} and $\mu_{ij'}$.*

(b) *For the Rado observable, actual preparation parameters do not witness a tree property for a fixed quantifier-free formula in one probe variable: for every $n \geq 1$ and every continuous connective $u\colon [0,1]^n \to \mathbb{R}$, no infinite sequence of instances of the formula*

$$u\big(R(y_1, x), \ldots, R(y_n, x)\big),$$

with actual mixtures as parameters, is uniformly k-inconsistent, for any integer $k \geq 2$.

The idea of (a) is that a single mixture over a coset tests many checks at once. In the language of error correction, the labels g_{im} are checks and Bob's label b is an error. If the error commutes with the checks g_{im}, $m < j$, then no string of the subgroup $\operatorname{span}(g_{im} : m < j)$ anticommutes with it, and either every string of the coset $g_{ij} + \operatorname{span}(g_{im} : m < j)$ does or none does, according as the error anticommutes with g_{ij} or not; if it anticommutes with one of the checks g_{im}, $m < j$, then exactly half of the strings of the subgroup, and exactly half

of those of the coset, anticommute with it. For $j < j'$, the pair of responses to μ_{ij} and $\mu_{ij'}$ at a pure probe is one of

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

These cases correspond, respectively, to the first anticommuting check occurring at j, before j, between j and j', at j', or not yet occurring. At a mixed probe the pair (x, y) lies in their convex hull, on which $x/2 + y \leq 1$ (see Figure 12); if both coordinates exceeded $2/3$, this inequality would fail. The bound is sharp: mixing a pure probe with pair $(1, \frac{1}{2})$ with weight $2/3$ and one with pair $(0, 1)$ with weight $1/3$ gives $(2/3, 2/3)$. No error has two first anticommuting checks in one row, and finitely many conditions on different rows, which concern disjoint families of linearly independent checks, can be satisfied together. For an entire path, the prescribed values extend to a linear functional on V, so the corresponding pure probe in V^* realizes it; this need not be an error of finite support. On pure probes, the difference between the responses to the uniform mixtures on the coset and on the subgroup is the indicator of a syndrome sector, the set of errors whose syndrome on the checks $g_{i0}, \ldots, g_{ij}$ is $(0, \ldots, 0, 1)$. This identity is the counterpart, for responses, of a familiar formula of error correction: for a stabilizer group $\mathcal{G}$ and a character χ of $\mathcal{G}$, the projector onto the joint eigenspace of the elements g of $\mathcal{G}$ with eigenvalues $\chi(g)$ is the average $|\mathcal{G}|^{-1} \sum_{g \in \mathcal{G}} \chi(g)\, g$. The identity, however, requires the checks only to be linearly independent, not to commute. Each label a of Alice defines the character $f \mapsto (-1)^{f(a)}$ of the compact group V^*, and combinations of averages over cosets recover, by Fourier inversion on V^*, the indicators of cylinder sets, the sets of functionals with prescribed values at finitely many labels. The idea of (b) is that the responses of the Rado observable depend only on the marginals of a probe, the probabilities q_a with which it answers 1 at the vertices a, and every vector of marginals is attained. Given an infinite sequence of conditions on marginals, a gliding hump argument, familiar from Banach space theory, yields a subsequence whose parameters place their weight, up to small errors, on a common finite set of vertices and on pairwise disjoint blocks, and one probe built block by block nearly satisfies any k of these conditions.

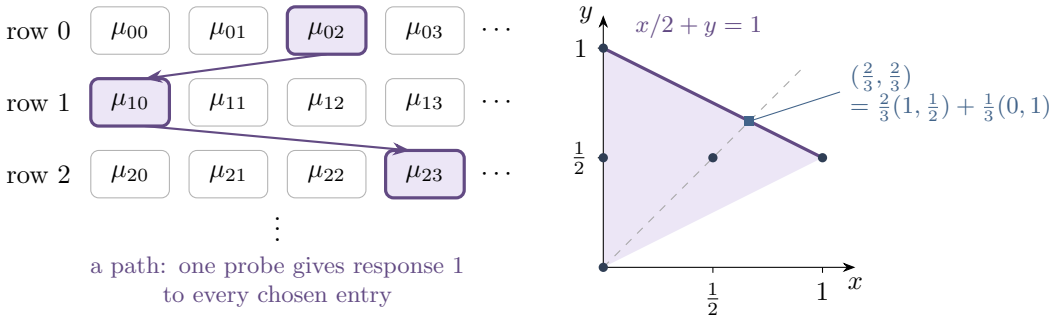


FIGURE 12. The array of coset mixtures, and the convex hull of the pairs of responses to two entries of one row.

For the physicist or operator theorist. The two observables differ in their probe algebras: for the commutation observable the probe algebra is the algebra of continuous functions on the compact group V^*, in which the differences between the averages over a coset of labels and over its subgroup are indicators of cylinder sets, whereas for the Rado observable the responses of mixtures depend only on the marginals of a probe.

For the logician. Alice's preparation μ_{ij} is the random check obtained by multiplying g_{ij} by a uniformly random product of the earlier checks g_{im}, $m < j$, of row i; against a fixed error, a label of Bob, it answers 1 with certainty if g_{ij} is the first of the checks $g_{i0}, \dots, g_{ij}$ that anticommutes with the error, with probability $\frac{1}{2}$ if an earlier check anticommutes with it, and with probability 0 otherwise. The uniform mixture on the subgroup $\mathrm{span}(g_{im} : m < j)$ answers 1 with probability $\frac{1}{2}$ if an earlier check anticommutes with the error and with probability 0 otherwise; consequently, the difference between the response to μ_{ij} and the response to this mixture is the indicator of a syndrome sector, the set of errors whose first anticommuting check in row i is g_{ij}. For the Rado observable, the fixed quantifier-free formulas of part (b) admit no witnessing array all of whose preparation parameters are actual mixtures. The construction of Theorem 28.1 supplies such a witness in an elementary extension by taking limits of finite uniform block mixtures.

The distinction between actual and idealized configurations, which we drew for ladders in Part I, thus reappears for trees.

VARIATIO 30

a 1 Clav. Quodlibet

The commutation relation

In this variation we follow a single relation of quantum mechanics through the three parts: a momentum kick and a translation commute up to a phase. Read with bounded kicks, with unbounded kicks and with binary kicks, it yields an example for each part (see Proposition 30.1 below).

For a particle on a line, with wave functions $\psi \in L^2(\mathbb{R})$ of the position q, let T_x be the translation by a real number x, given by $(T_x\psi)(q) = \psi(q-x)$, and let K_k be the momentum kick by a real number k, given by $(K_k\psi)(q) = e^{ikq}\psi(q)$. With the position operator Q, multiplication by q, and the momentum operator $P = -i\,d/dq$, in units in which the reduced Planck constant $\hbar$ equals 1, one has $T_x = e^{-ixP}$ and $K_k = e^{ikQ}$, and $K_k^* P K_k = P + k$: the kick shifts momentum by k. The two families satisfy

$$K_k T_x = e^{ikx}\, T_x K_k,$$

the Weyl commutation relation, the bounded form of the canonical commutation relation $QP - PQ = iI$, which holds on a dense domain; by the Stone–von Neumann theorem, it determines (Q, P) up to unitary equivalence for irreducible, strongly continuous families. On a qubit, the bit flip X, which exchanges the basis vectors $|0\rangle$ and $|1\rangle$, is a translation, the phase flip Z, which multiplies $|s\rangle$ by $(-1)^s$ for $s \in \{0, 1\}$, is a kick, and the two anticommute, $ZX = -XZ$: the Pauli group, generated by these translations and kicks on each qubit together with the scalar iI, is the Weyl–Heisenberg group over the two-element field, with the phase $(-1)^{\omega(v,w)}$ of the commutation relation of two Pauli strings with labels v and w (see Variatio 22) in place of e^{ikx}. The phase e^{ikx} also determines the intensity that Bob observes when Alice's photon, prepared with wavenumber k, reaches his detector at position x: the interference fringe of Variatio 16 is $\frac{1}{2}(1+\mathrm{Re}\, e^{ikx})$. Read through bipartite observables in which Alice chooses a kick k and Bob a translation x, the position of his detector, the commutation phase yields three observables according to the range of its parameters, one for each of the three dividing lines. For a sequence $(k_n)_{n\in\mathbb{N}}$ of wavenumbers, recall from Variatio 14 the fringe observable $C = \sum_n |e_n\rangle\langle e_n| \otimes M_{f_n}$ on $\ell^2(\mathbb{N}) \otimes L^2[0, 2\pi]$, where $(e_n)_{n\in\mathbb{N}}$ is the standard basis of $\ell^2(\mathbb{N})$ and M_{f_n} is the operator of multiplication by $f_n(x) = \frac{1}{2}(1 + \cos k_n x)$. Over the two-element field, on countably many qubits, the commutation relation becomes the commutation form of the structure (V, ω) of labels of Pauli strings, and the observable that reads it is the commutation observable of Variatio 28.

Proposition 30.1.

(a) Bounded parameters: stable. *If Alice's wavenumbers k_n, for $n \in \mathbb{N}$, lie in a bounded set, then the fringe observable C is stable.*

(b) Lacunary parameters: not NIP. *If $k_n = 8^n$ for $n \in \mathbb{N}$, then the fringe observable C is not NIP, and its idealized responses contain a copy of $\beta\mathbb{N}$.*

(c) Binary parameters: simple, and not simple once mixed. *The theory of* (V, ω) *is simple and unstable, and the commutation observable has the tree property of the second kind, witnessed by actual mixtures of Alice's preparations.*

Proof. (a) Because C is diagonal in Alice's factor, the off-diagonal parts of her states do not contribute, and the effective observable of a state ρ is multiplication by $\sum_n \langle e_n, \rho e_n \rangle f_n$. The functions f_n are uniformly bounded and, since $|f_n'(x)| \leq \frac{1}{2}|k_n|$, equicontinuous; by the Arzelà–Ascoli theorem they form a relatively compact set in $C[0, 2\pi]$, and so does its closed convex hull. A relatively norm-compact set of effective observables is relatively weakly compact, and the observable is stable, by Theorem 5.1(A2). (b) This is Proposition P.45. (c) This is Theorem 22.1 together with Theorem 29.1(a). □

The proposition concerns three restrictions of one relation, not one observable in three classes: the bounded and the unbounded fringes are two observables, and the qubit version is a third. With bounded parameters, the relation satisfies Grothendieck's double limit condition; with the wavenumbers 8^n it is not NIP. Over the two-element field it is the arithmetic of error correction, simple in itself and not simple once preparations are mixed.

For the physicist or operator theorist. In operator terms, the three regimes are three behaviors of the effective observables of Alice: with bounded wavenumbers, the operators of multiplication by averages of the fringes f_n form a relatively norm-compact set; with the wavenumbers 8^n, some sequence of these operators is equivalent, in the operator norm, to the unit vector basis of ℓ^1 (see Variatio 11); and over the two-element field the differences between the averages over a coset of checks and over its subgroup are indicators of syndrome sectors.

For the logician. A translation T_x moves a particle by x and a kick K_k adds k to its momentum; applied in the two orders, they produce the same state up to the factor e^{ikx}, which is invisible for a single state but observable in interference, where two amplitudes are added. In the fringe observable this phase is measured directly: a photon of wavenumber k is detected at the position x with relative intensity $\frac{1}{2}(1 + \cos kx)$, and Bob's choice of x is the choice of a translation. On a qubit the bit flip X is the translation of $\{0, 1\}$ by 1 and the phase flip Z is a kick, so the relation $ZX = -XZ$ is the same relation with phases ± 1; on n qubits it becomes the commutation relation ω of Pauli strings, which a stabilizer measurement reads. The proposition follows one physical relation through the three dividing lines, according to whether the parameters range over a bounded set of reals, over the lacunary set $\{8^n\}$, or over the two-element field.

Aria da capo e fine

Physics lives on idealized limits. We modeled idealized preparations by limiting states of actual preparations, and represented their sequences by elements of ultrapowers. For a single system, every bounded expectation passed to the limit; in the variations, the question was what happened when two systems were combined.

Difficulties begin with two systems, and there they stayed. A simple example, the staircase, was decisive from the start: whichever of the two oscillators, Alice's and Bob's, was driven first to infinite energy outran the other, and whether Bob's had the higher energy depended on which was idealized first. The comparison of the two energies is therefore unstable, and in Part II it proved to be NIP.

The agreement of the two orders had a finite approximation consequence. Where the orders agreed, every state of Alice's algebra was simulated in its responses by finite mixtures of actual preparations, uniformly over Bob's probes and with his system idealized first. Their common values gave a symmetric independent product on the space of stable observables; at each stable observable, the expectation of the product was the unique separately weak* continuous extension from pairs of actual states. For uniformly stable observables, the number of transcripts of adaptive expectation queries was bounded by a polynomial at each positive margin, and their labels could be predicted in a fixed-threshold game with a finite mistake bound uniform in the threshold. Where uniform stability failed, the number of transcripts grew exponentially at some margin, even for observables that remained stable.

The theme of this memoir is Grothendieck's theorem, which says exactly when the two orders agree, and it lay beneath the variations as their bass. In Part I it was heard plainly. In Part II it was broken, and, for the observables that are NIP, what remained of it was sequential: every idealized response was still a limit of a sequence of responses of actual preparations. In Part III it fell silent.

A logician recognizes the staircase at once: it is an instance of the order property, and in the variations the three dividing lines that Shelah drew became three distinctions among measurements. The first, stability, is drawn by Grothendieck's double limit condition itself; the second, NIP, separates the failures of the condition in which idealized responses remain such limits from those in which they do not; and the third, simplicity, is drawn not by limits but by the amalgamation of independent types, through the Kim–Pillay theorem. Mixing, the operation that distinguishes quantum observables from arbitrary formulas, did more than simulate idealized preparations: for the observables that are NIP, it made hereditary separability of the idealized responses imply their metrizability, and, for the observables that, like the comparison of energies, read a relation between the vectors of two orthonormal bases, it made simplicity imply NIP.

In Part I we studied the observables that satisfy Grothendieck's double limit condition, in Part II the observables that are NIP, and in Part III the amalgamation of independent types, which in simple theories takes the place of the uniqueness of nonforking extensions.

Proofs

We collect here the proofs too long for the variations, in the order of the variations whose results they prove. They use the notation of the variations, which we recall where needed.

P.1. Part I: stability and uniform stability

Throughout this subsection, C is a bipartite observable on $H_1 \otimes H_2$. For a Hilbert space H, let $\mathfrak{T}(H)$ be the Banach space of trace-class operators on H with the trace norm $\|\tau\|_1 = \operatorname{Tr}|\tau|$, where $|\tau| = (\tau^*\tau)^{1/2}$; the dual of $\mathfrak{T}(H)$ is $B(H)$, each operator A acting by $\tau \mapsto \operatorname{Tr}(\tau A)$, and the bidual of $\mathfrak{T}(H)$ is $B(H)^*$, which contains the states of $B(H)$. The *slice maps* of C are the bounded operators $T_C \colon \mathfrak{T}(H_1) \to B(H_2)$ and $S_C \colon \mathfrak{T}(H_2) \to B(H_1)$ determined by

$$\operatorname{Tr}\bigl(\tau' T_C(\tau)\bigr) = \operatorname{Tr}\bigl((\tau \otimes \tau')C\bigr) = \operatorname{Tr}\bigl(\tau S_C(\tau')\bigr) \qquad \text{for all } \tau \in \mathfrak{T}(H_1) \text{ and } \tau' \in \mathfrak{T}(H_2);$$

on density operators, T_C is the effective map of Variatio 1, and both slice maps have norm at most $\|C\|$. For a state ω_1 of $B(H_1)$, the operator Z_{ω_1} of Variatio 1 satisfies $\operatorname{Tr}(\sigma Z_{\omega_1}) = \omega_1(S_C(\sigma))$ for every density operator σ on H_2, so Z_{ω_1} is the image of ω_1 under the transpose $S_C' \colon B(H_1)^* \to B(H_2)$ of S_C. Symmetrically, a state ω_2 of $B(H_2)$ leaves on Alice's side the operator $W_{\omega_2} = T_C'(\omega_2) \in B(H_1)$, which satisfies $\operatorname{Tr}(\tau W_{\omega_2}) = \omega_2(T_C(\tau))$ for every $\tau \in \mathfrak{T}(H_1)$. Computing the inner limits in the definitions of Variatio 1 gives

$$(\omega_1 \otimes_{\mathsf{A}} \omega_2)(C) = \omega_2(Z_{\omega_1}), \qquad (\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \omega_1(W_{\omega_2}) = \bigl\langle T_C''(\omega_1), \omega_2 \bigr\rangle, \tag{1}$$

where $T_C'' \colon B(H_1)^* \to B(H_2)^{**}$ is the bitranspose of T_C. In particular, the product with Alice's system idealized first is weak* continuous in Bob's state, and the product with Bob's system idealized first is weak* continuous in Alice's.

Fact P.1 (see [KR83, Section 4.3]). *Let H be a separable Hilbert space. The actual states are weak* dense in the set of states of $B(H)$, and every $\omega \in B(H)^*$ is a combination $c_1\omega_1 - c_2\omega_2 + ic_3\omega_3 - ic_4\omega_4$ of four states $\omega_1, \dots, \omega_4$ of $B(H)$ with $0 \le c_j \le \|\omega\|$ for $j = 1, \dots, 4$.*

Proof of Theorem 5.1. The equivalence of (A1) and (A4) is Theorem 3.1.

(A2) and (A4). Let $\mathcal{K} = T_C(\mathfrak{D}(H_1))$. Every τ in the unit ball of $\mathfrak{T}(H_1)$ is a combination $c_1\rho_1 - c_2\rho_2 + ic_3\rho_3 - ic_4\rho_4$ of density operators ρ_j with $0 \le c_j \le 1$, given by the positive and negative parts of the Hermitian parts of τ; the slice map T_C therefore sends the unit ball into the sum of four copies of the set $\{cA : c \in \mathbb{C},\ |c| \le 1,\ A \in \mathcal{K}\}$, which is relatively weakly compact when $\mathcal{K}$ is. Hence (A2) holds exactly when T_C is weakly compact. By Gantmacher's theorem [Gan40], this happens exactly when T_C'' maps $B(H_1)^*$ into $B(H_2)$, regarded as a subspace of $B(H_2)^{**}$. For a state ω_1 of $B(H_1)$ and a density operator σ on H_2, the transpose T_C' sends σ to $S_C(\sigma)$, and therefore $\langle T_C''(\omega_1), \sigma\rangle = \omega_1(S_C(\sigma)) = \operatorname{Tr}(\sigma Z_{\omega_1})$. Since the actual states determine an operator, the functional $T_C''(\omega_1)$ belongs to $B(H_2)$ exactly when $T_C''(\omega_1) = Z_{\omega_1}$. By (1) and Fact P.1, by which the states span $B(H_2)^*$, this happens exactly when $(\omega_1 \otimes_{\mathsf{B}} \omega_2)(C) = \omega_2(Z_{\omega_1}) = (\omega_1 \otimes_{\mathsf{A}} \omega_2)(C)$ for every state ω_2 of $B(H_2)$. The states also span $B(H_1)^*$, so T_C'' maps $B(H_1)^*$ into $B(H_2)$ exactly when (A4) holds.

(A2) and (A3). If T_C is weakly compact, the Davis–Figiel–Johnson–Pełczyński theorem [**DFJP74**] gives a reflexive Banach space X and bounded operators $a\colon \mathfrak{T}(H_1) \to X$ and $J\colon X \to B(H_2)$ with $T_C = J \circ a$. Let u be the restriction of a to $\mathfrak{D}(H_1)$, and let $v(\sigma)$ be the functional $\xi \mapsto \operatorname{Tr}(\sigma J(\xi))$ on X, for $\sigma \in \mathfrak{D}(H_2)$. Both maps are bounded and affine, and $v(\sigma)(u(\rho)) = \operatorname{Tr}(\sigma\, T_C(\rho)) = \varphi_C(\rho, \sigma)$. Conversely, suppose that bounded maps u and v, affine or not, factor the expectation through the pairing of a reflexive space X with X^*, and let $(\rho_k)_{k\in\mathbb{N}}$ and $(\sigma_l)_{l\in\mathbb{N}}$ be sequences along which both iterated limits of $\varphi_C(\rho_k, \sigma_l)$ exist. Bounded sequences in a reflexive space have weakly convergent subsequences, by the Eberlein–Šmulian theorem [**Ebe47**, **Š40**]; pass to subsequences along which $u(\rho_k)$ converges weakly to $\xi \in X$ and $v(\sigma_l)$ converges weakly to $\xi' \in X^*$. The iterated limit taken first over l is $\lim_k \xi'(u(\rho_k)) = \xi'(\xi)$, and the one taken first over k is $\lim_l v(\sigma_l)(\xi) = \xi'(\xi)$, because X is the dual of X^*. Hence (A3) implies (A1), which is equivalent to (A2).

(A5), (A6) and (A7). The equivalence of (A4) and (A5) is Theorem 9.1(a), and that of (A4) and (A6) is Theorem 10.1. For (A7), recall from Variatio 1 that, for sequences $(\rho_n)_{n\in\mathbb{N}}$ and $(\sigma_m)_{m\in\mathbb{N}}$ of density operators and nonprincipal ultrafilters $\mathcal{U}$ and $\mathcal{V}$ on $\mathbb{N}$, the two orders of realizing the classes $[\rho_n]_\mathcal{U}$ and $[\sigma_m]_\mathcal{V}$ in an iterated ultrapower give the two products of the states $\lim_{n,\mathcal{U}} \rho_n$ and $\lim_{m,\mathcal{V}} \sigma_m$ at C, and that every state is the limit of a sequence of density operators along some ultrafilter. Condition (A7) is therefore (A4), and the common value is the independent product of Theorem 6.4. □

We now turn to uniform stability. We record the orientation of a ladder or of a tree by a sign. A ladder at margin δ with threshold t has *sign* $\theta \in \{1, -1\}$ if $\theta(\varphi_C(\rho_i, \sigma_j) - t) \geq \delta/2$ for $i \leq j$ and $\theta(\varphi_C(\rho_i, \sigma_j) - t) \leq -\delta/2$ for $i > j$; the ladders of Variatio 2 are the ladders of either sign. For an integer $n \geq 0$, we write $\{0,1\}^n$ for the set of 0–1 sequences of length n and $\{0,1\}^{<n}$ for the set of those of length less than n, and, for sequences s and η, we write $s \sqsubseteq \eta$ if s is an initial segment of η and $s^\frown\iota$ for s followed by the digit ι. The nodes of the complete binary tree of height n are the sequences $s \in \{0,1\}^{<n}$, the branches are the sequences $\eta \in \{0,1\}^n$, and a branch η passing through s turns right at s if $s^\frown 1 \sqsubseteq \eta$ and left if $s^\frown 0 \sqsubseteq \eta$. A tree of Variatio 7, with probes σ_s and preparations ρ_η, has *sign* θ if $\theta(\varphi_C(\rho_\eta, \sigma_s) - t) \geq \delta/2$ whenever η turns right at s and $\theta(\varphi_C(\rho_\eta, \sigma_s) - t) \leq -\delta/2$ whenever η turns left at s. For $\delta > 0$, let $d_C(\delta) \in \{0, 1, \dots, \infty\}$ be the supremum of the heights of the trees of C at margin δ, over all thresholds and signs.

Lemma P.2. *Let $\delta > 0$, and let $h \geq 1$ be an integer. Every ladder of C at margin δ with 2^h rungs, threshold t and sign θ contains a tree of C of height h at margin δ, with the same threshold and sign.*

Proof. This is the passage from ladders to trees of stability theory [**She90**, Chapter II]. The probe $\sigma_{2^{h-1}}$ of the middle rung separates the preparations of the first half of the ladder from those of the second, and each half is a ladder with 2^{h-1} rungs, the same threshold and the same sign; induction on h places the trees given by the two halves above this probe. □

For a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the expectation of the ultrapower $\mathcal{M}_C^\mathcal{U}$ of Variatio 1 is the function $\Phi_C^\mathcal{U}(x, y) = \lim_{k,\mathcal{U}} \varphi_C(x_k, y_k)$ of the classes of sequences $x = (x_k)_{k\in\mathbb{N}}$ of Alice's density operators and $y = (y_k)_{k\in\mathbb{N}}$ of Bob's; ladders and trees for $\Phi_C^\mathcal{U}$ are defined as for φ_C.

First part of the proof of Theorem 7.1. We show that conditions (B1), (B2) and (B3) are equivalent. If the trees of C have bounded heights at every margin, so do the ladders of C, by Lemma P.2; hence (B2) implies (B1).

(B1) implies (B3). Suppose that $\Phi_C^\mathcal{U}$ is unstable for a nonprincipal ultrafilter $\mathcal{U}$. The proof of Theorem 2.1 uses only the values of the function and applies to every bounded real function

on a product of two sets; for $\Phi_C^{\mathcal{U}}$ the proof provides, for some $\delta > 0$, an infinite ladder for $\Phi_C^{\mathcal{U}}$ at margin δ, formed by sequences $x_i = (x_{i,k})_{k\in\mathbb{N}}$ and $y_j = (y_{j,k})_{k\in\mathbb{N}}$, with a threshold t and a sign θ. Fix a positive integer N. For each of the N^2 pairs (i, j) with $1 \le i, j \le N$, the set of indices k with $\theta(\varphi_C(x_{i,k}, y_{j,k}) - t) \ge \delta/4$, if $i \le j$, or $\le -\delta/4$, if $i > j$, belongs to $\mathcal{U}$. For k in the intersection of these N^2 sets, the density operators $x_{i,k}$ and $y_{j,k}$ form a ladder of C of length N at margin $\delta/2$. Hence the ladders of C at margin $\delta/2$ have unbounded length, and C is not uniformly stable.

(B3) for one ultrafilter implies (B2). Suppose that, for some $\delta > 0$, the observable C has trees $(\rho_\eta^{(n)}, \sigma_s^{(n)})$ of every height $n \ge 1$ at margin δ, with thresholds t_n and signs θ_n, and let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$. On a set of indices n that belongs to $\mathcal{U}$, the sign θ_n is constant, equal to θ; each root has preparations on both sides of the threshold t_n, so $|t_n| \le \|C\|$, and we let $t = \lim_{n,\mathcal{U}} t_n$. For an infinite 0–1 sequence η, let x_η be the sequence $(\rho_{\eta|n}^{(n)})_{n\ge1}$, where $\eta|n$ is the initial segment of length n, and, for a finite 0–1 sequence s, let y_s be the sequence $(\sigma_s^{(n)})_{n\ge1}$, with arbitrary density operators for $n \le |s|$, where $|s|$ is the length of s. If η turns right at s, then so does $\eta|n$ for every $n > |s|$, and passing to the limit along $\mathcal{U}$ gives $\theta(\Phi_C^{\mathcal{U}}(x_\eta, y_s) - t) \ge \delta/2$; if η turns left at s, the opposite inequality holds. The classes of these sequences therefore form an infinite tree for $\Phi_C^{\mathcal{U}}$ at margin δ. Suppose now that $\Phi_C^{\mathcal{U}}$ is stable. After the affine rescaling $c(\Phi_C^{\mathcal{U}} + \|C\|)$, with $c = 1/(2\|C\|)$, which takes values in $[0, 1]$ and multiplies gaps by c, Iovino's theorem [**Iov99b**, Theorem 3.3] gives a reflexive Banach space X and maps u and v, from the two sorts of the ultrapower into the unit balls of X and of X^*, with $c(\Phi_C^{\mathcal{U}} + \|C\|) = v(\cdot)(u(\cdot))$. Let Y be the closed span in X^* of the countably many vectors $v(y_s)$; the space Y is separable and reflexive, so the dual Y^* is separable. If two infinite sequences $\eta \ne \eta'$ first differ immediately after the segment s, then $|v(y_s)(u(x_\eta) - u(x_{\eta'}))| \ge c\delta$, and the restrictions to Y of the vectors $u(x_\eta)$ are therefore $2^{\aleph_0}$ points of Y^* at mutual distances at least $c\delta$, which is impossible in a separable space. Hence, if $\Phi_C^{\mathcal{U}}$ is stable for one nonprincipal ultrafilter $\mathcal{U}$, the trees of C have bounded heights at every margin.

Since (B3) for every ultrafilter implies (B3) for one, the three conditions are equivalent. □

Every uniformly stable observable is stable, since, by Theorem 2.1, an unstable observable has an infinite ladder at some margin. The sector comparison of Example 7.2 shows that the converse fails.

Proposition P.3. *Every self-adjoint element of the minimal tensor product $B(H_1) \otimes_{\min} B(H_2)$ is uniformly stable, and so is every bipartite observable whose expectation factors through a Hilbert space by bounded maps, as in Theorem 4.2.*

Proof. For a finite sum $C' = \sum_{j=1}^{r} A_j \otimes B_j$ of products, the expectation of the ultrapower is $\Phi_{C'}^{\mathcal{U}}(x, y) = \sum_j \lim_{k,\mathcal{U}} \operatorname{Tr}(x_k A_j) \cdot \lim_{k,\mathcal{U}} \operatorname{Tr}(y_k B_j)$, a sum of products of a function of x and a function of y, and $|\Phi_C^{\mathcal{U}} - \Phi_{C'}^{\mathcal{U}}| \le \|C - C'\|$; the argument of Theorem 4.3(a) shows that $\Phi_C^{\mathcal{U}}$ is stable when C lies in the norm closure of such sums, and C is uniformly stable by (B3) of Theorem 7.1. In the Hilbert case, regard the Hilbert space as a real Hilbert space with the inner product $\operatorname{Re}\langle\cdot,\cdot\rangle$, which gives the same values, since the expectation is real; the trees of C then have bounded heights at every margin by Theorem 8.1, which is (B2). □

For the counting in (B4), fix $\delta > 0$, a threshold t and a sign θ. For an integer $m \ge 0$, a *probe tree* of depth m is a family $(\sigma_s)_{s\in\{0,1\}^{<m}}$ of Bob's density operators, and a branch $\eta \in \{0, 1\}^m$ is *realized* by a density operator ρ of Alice if $\theta(\varphi_C(\rho, \sigma_s) - t) \ge \delta/2$ whenever η turns right at s and $\theta(\varphi_C(\rho, \sigma_s) - t) \le -\delta/2$ whenever η turns left at s. The realized branches

are the transcripts that m adaptive probes can return, and preparations that realize different branches are told apart by the probes. Let $g_C(m,\delta)$ be the supremum, over thresholds, signs and probe trees of depth m, of the number of realized branches, and, for an integer $d \geq 0$, let $\binom{m}{\leq d} = \sum_{i=0}^{d} \binom{m}{i}$, with $\binom{m}{\leq -1} = 0$.

Proposition P.4 (Sauer, Shelah; see [BDPSS09, RST15]). *Let $\delta > 0$ and let $m \geq 0$ be an integer. If $d = d_C(\delta)$ is finite, then $g_C(m,\delta) \leq \binom{m}{\leq d}$, a polynomial in m of degree d. If $d_C(\delta)$ is infinite, then $g_C(m,\delta) = 2^m$.*

The first statement is the tree form of the Sauer–Shelah lemma, applied to the class of the thresholded responses $\sigma \mapsto \varphi_C(\rho,\sigma)$, for $\rho \in \mathfrak{D}(H_1)$, whose trees, in the sense of these references, are the trees of C with threshold t and sign θ. The proof splits the preparations at the root probe and uses that the two parts cannot both carry trees of the largest height. The second statement holds because a tree of height at least m, restricted to the nodes of length less than m, realizes every branch of a probe tree of depth m.

For the prediction game of (B5), fix a threshold t and a margin $\delta > 0$. Alice prepares an unknown density operator ρ; in each round, Bob chooses a probe σ with $|\varphi_C(\rho,\sigma) - t| \geq \delta/2$, a learner predicts on which side of t the value lies, and is then told the answer. Let $\mathrm{mist}_C(t,\delta)$ be the least integer n such that some deterministic learner makes at most n mistakes for every ρ and every sequence of probes, or ∞ if there is none, and let $d_C(t,\delta)$ be the supremum of the heights of the trees at margin δ with threshold t and sign 1; reflecting a tree, that is, exchanging the two subtrees at every node, changes the sign of the tree, so $\sup_t d_C(t,\delta) = d_C(\delta)$.

Proposition P.5 (Littlestone [Lit88]). *For every real number t and every $\delta > 0$, we have $\mathrm{mist}_C(t,\delta) = d_C(t,\delta)$.*

Littlestone's proof applies without change: an adversary who walks down a tree, against the predictions of the learner, forces one mistake at each node, and the learner that predicts the side whose consistent preparations carry the higher tree errs at most $d_C(t,\delta)$ times, by the splitting argument behind Proposition P.4.

Second part of the proof of Theorem 7.1. By Proposition P.4, condition (B2) implies (B4), and if (B2) fails at a margin δ, then $g_C(m,\delta) = 2^m$ for every m, which is the last statement of the theorem and contradicts (B4). By Proposition P.5, condition (B5) is (B2) with the threshold fixed and the bound uniform in the threshold. □

The bound survives when the threshold varies from probe to probe. Let $M = \|C\|$, so that every response lies in $[-M,M]$, and choose $K = \lceil 4M/\delta \rceil + 1$ equally spaced thresholds $t_1,\dots,t_K$ in $[-M,M]$, including both endpoints; every point of $[-M,M]$ then lies within $\delta/4$ of one of them. Run a learner of Proposition P.5 at each t_j, with margin $\delta/2$, so that it makes at most $d_C(t_j,\delta/2) \leq d_C(\delta/2)$ mistakes. When a probe comes with a threshold $t \in [-M,M]$, consult and then update only the learner at a threshold t_j within $\delta/4$ of t: a response at least $t + \delta/2$ is at least $t_j + \delta/4$, and a response at most $t - \delta/2$ is at most $t_j - \delta/4$, so the probe is admissible for t_j at margin $\delta/2$, and its response lies on the same side of t_j as of t. A threshold outside $[-M,M]$ leaves only one possible answer, which the combined learner predicts. The combined learner therefore errs at most $K\, d_C(\delta/2)$ times, a number that is finite under (B2).

P.2. Part II: NIP

In this subsection we prove the results of Part II. Theorem 19.1 consists of Theorem P.29, Corollary P.31 and Theorem P.32. In Theorem 16.1, part (a) is Theorem P.40(a), part (b)

is Theorem P.40(b) together with Proposition P.41, part (c) is Theorem P.42, part (d) is Theorem P.44, and part (e) is Proposition P.45. Theorem 15.1 is Theorem P.50, the results on the Ising model are Theorem P.54, and Theorem 18.1 consists of Proposition P.56, Proposition P.57 and Theorem P.58. The classical results on Rosenthal compacta, the dichotomies of Rosenthal and of Bourgain, Fremlin and Talagrand, Todorčević's trichotomy and the heptachotomy of Argyros, Dodos and Kanellopoulos, are recalled as facts, with references.

A family of continuous functions is NIP exactly when the closure consists of Baire-1 functions. We begin with the topological dichotomy on which Part II rests: a pointwise bounded family of continuous functions on a compact metrizable space either has the independence property or has a pointwise closure inside the functions of the first Baire class. In what follows, the functions are responses of Alice's preparations and the points are Bob's probes.

Functions of the first Baire class. Throughout, the set $\mathbb{N}$ of natural numbers contains 0. For a topological space Y, let $C_p(Y)$ be the set of continuous functions $Y \to \mathbb{R}$ with the topology of pointwise convergence. A function $f\colon Y \to \mathbb{R}$ is of the *first Baire class*, or *Baire-1*, if f is the pointwise limit of a sequence of continuous functions, and we write $B_1(Y)$ for the set of these functions, again with the topology of pointwise convergence. For $\mathcal{R} \subseteq \mathbb{R}^Y$, we write $\overline{\mathcal{R}}$ for the closure of $\mathcal{R}$ in $\mathbb{R}^Y$, and we call $\mathcal{R}$ *pointwise bounded* if $\sup_{r\in\mathcal{R}} |r(y)| < \infty$ for every $y \in Y$.

We use the classical characterizations of Baire-1 functions on Polish spaces, the separable, completely metrizable spaces: condition (3) of Fact P.6 below is Baire's characterization, and condition (2) concerns F_σ sets, the countable unions of closed sets.

Fact P.6 (see [Tod97, Section 10]). *Let Y be a Polish space and $f\colon Y \to \mathbb{R}$ a function. The following are equivalent:*

(1) *the function f is Baire-1;*
(2) *for every open $U \subseteq \mathbb{R}$, the set $f^{-1}[U]$ is F_σ;*
(3) *for every nonempty closed $F \subseteq Y$, the restriction $f|_F$ has a point of continuity.*

Moreover, if f is not Baire-1, then there are nonempty countable sets $Z_0, Z_1 \subseteq Y$ with $\overline{Z_0} = \overline{Z_1}$ and reals $a < b$ such that $f \le a$ on Z_0 and $f \ge b$ on Z_1.

The last clause of Fact P.6 lets us test membership in $B_1(Y)$ on countable sets, which is what makes pointwise closures manageable.

Lemma P.7 (Bourgain, Fremlin, Talagrand [BFT78]). *Let Y be a Polish space and $\mathcal{R} \subseteq B_1(Y)$ a pointwise bounded family. The following are equivalent:*

(i) *the family $\mathcal{R}$ is relatively compact in $B_1(Y)$;*
(ii) *every sequence in $\mathcal{R}$ has a cluster point in $B_1(Y)$;*
(iii) *we have $\overline{\mathcal{R}} \subseteq B_1(Y)$.*

The independence property. The obstruction to Baire-1 closures is combinatorial, as Theorem P.11 below shows: the obstruction is a sequence of functions whose sublevel and superlevel sets meet in every finite pattern.

Definition P.8. Let Y be a set. A family $\mathcal{R} \subseteq \mathbb{R}^Y$ has the *independence property* if there are a sequence $(r_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ and reals $a < b$ such that, for all disjoint finite sets $E, F \subseteq \mathbb{N}$, there is $y \in Y$ with $r_n(y) \le a$ for $n \in E$ and $r_n(y) \ge b$ for $n \in F$. Otherwise, the family $\mathcal{R}$ is *NIP*.

On the Cantor space $\{0,1\}^{\mathbb{N}}$ of infinite sequences of zeros and ones, the family of coordinate functions p_n, given by $p_n(y) = y(n)$ for $n \in \mathbb{N}$, has the independence property, and the closure of $\{p_n : n \in \mathbb{N}\}$ in $\mathbb{R}^{\{0,1\}^{\mathbb{N}}}$ is a copy of $\beta\mathbb{N}$, the space of ultrafilters on $\mathbb{N}$. The next two facts say that this example is the only obstruction.

Fact P.9 (Rosenthal's dichotomy; see [Ros74, Tod97]). *Let Y be a Polish space and $(r_n)_{n\in\mathbb{N}}$ a pointwise bounded sequence in $C_p(Y)$. Then the sequence (r_n) has either a pointwise convergent subsequence or a subsequence whose closure in $\mathbb{R}^Y$ is homeomorphic to $\beta\mathbb{N}$.*

Fact P.10 (Bourgain–Fremlin–Talagrand; see [BFT78, Theorem 4G]). *Let Y be a Polish space and $\mathcal{R} \subseteq C_p(Y)$ a pointwise bounded family. Then $\overline{\mathcal{R}} \subseteq B_1(Y)$ if and only if, for every sequence $(r_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ and all reals $a < b$, there is $I \subseteq \mathbb{N}$ such that no $y \in Y$ satisfies $r_n(y) \leq a$ for all $n \in I$ and $r_n(y) \geq b$ for all $n \notin I$.*

When the space Y is compact, the condition in Fact P.10 reduces to a statement about finite patterns, and we obtain the dichotomy announced at the beginning of this section.

Theorem P.11. *Let Y be a compact metrizable space and $\mathcal{R} \subseteq C_p(Y)$ a pointwise bounded family. Then exactly one of the following holds:*

(i) *we have $\overline{\mathcal{R}} \subseteq B_1(Y)$;*
(ii) *the family $\mathcal{R}$ has the independence property.*

Proof. By compactness of Y, the condition of Fact P.10 fails for a sequence (r_n) and reals $a < b$ exactly when every finite pattern is realized [DIMW+26, Section 1]. □

We shall also need to replace a space of probes by a dense subset, and the replacement does not affect the independence property.

Lemma P.12. *Let Y be a topological space, $\mathcal{R} \subseteq C_p(Y)$ a family and $L \subseteq Y$ a dense set. Then the family $\mathcal{R}$ is NIP if and only if the family $\{r|_L : r \in \mathcal{R}\}$ of restrictions is NIP.*

Sketch. If the restrictions have the independence property, then so does $\mathcal{R}$. Conversely, if (r_n) and $a < b$ witness the independence property of $\mathcal{R}$, and $a < a' < b' < b$, then each finite pattern holds strictly, with a' and b', on a nonempty open set, which meets L; hence the restrictions witness the independence property with the reals $a' < b'$. □

Remark P.13 (Symmetry). For a formula of a first-order theory, NIP is symmetric in the two variables: the formula $\varphi(x, y)$ is NIP if and only if the formula φ^{opp}, given by $\varphi^{\mathrm{opp}}(y, x) = \varphi(x, y)$, is NIP (see [She90, Chapter II, Section 4] and [Sim15a, Lemma 2.5]). Assouad's bound makes the symmetry quantitative. The Vapnik–Chervonenkis (VC) dimension of φ is the largest size of a finite set of elements every subset of which is cut out by some instance $\varphi(x, c)$ of the formula, for a parameter c that depends on the subset; if φ has finite VC dimension d, then the formula φ^{opp} has VC dimension less than 2^{d+1} [Ass83]. Assouad's argument is finite and combinatorial, and applies, with the reals $a < b$ fixed, to real-valued functions. For uniform NIP, which bounds, for each gap $b - a$, the size of the finite sets of functions on which every pattern of values at most a and at least b is realized at some point (see Remark P.25 below), the roles of Alice and Bob are therefore interchangeable. The notion of Definition P.8, which asks for an infinite sequence inside a single family, is not symmetric: the relation $n \in u$ between numbers n and finite sets $u \subseteq \mathbb{N}$ has the independence property in the variable n but not in the variable u [KP18].

Rosenthal compacta. A *Rosenthal compactum* is a compact space homeomorphic to a compact subset of $B_1(Y)$, with the topology of pointwise convergence, for some Polish space Y. By Theorem P.11 and Lemma P.7, if Y is compact metrizable and $\mathcal{R} \subseteq C_p(Y)$ is pointwise bounded and NIP, then the closure $\overline{\mathcal{R}}$ is a Rosenthal compactum. The classification of Rosenthal compacta (see section P.2) will therefore classify the NIP families that arise below.

Separable Rosenthal compacta fall into finitely many canonical classes. The Baire-1 characterization reduces the analysis of NIP families to the topology of their pointwise closures.

For a countable, pointwise bounded NIP family of continuous functions on a compact metrizable space, this closure is a separable Rosenthal compactum. Todorčević's trichotomy distinguishes the nonmetrizable cases by hereditary separability and first countability; the canonical form of the trichotomy identifies minimal families whose closures embed into the corresponding compacta. We state these results together with the Argyros–Dodos–Kanellopoulos heptachotomy, following their notation [**ADK08**], and formulate the resulting levels of NIP for use with response families.

Three topological invariants. Recall that a space is hereditarily separable if each subspace has a countable dense subset, and first countable if each point has a countable base of neighborhoods. Every compact metrizable space is hereditarily separable, and by a theorem of Pol every hereditarily separable Rosenthal compactum is first countable (see [**Deb13**, Section 10]). For a Rosenthal compactum K we therefore have:

$$K \text{ metrizable} \implies K \text{ hereditarily separable} \implies K \text{ first countable}.$$

For separable Rosenthal compacta, the failure of each converse is witnessed by a canonical subspace, and these canonical subspaces are pointwise closures of families indexed by the dyadic tree, the set of finite sequences of zeros and ones (see Fact P.16 below).

The canonical compacta. The minimal families encode localization, one-sided limits and paired observations by functions on the dyadic tree and on the boundary of the tree. Let $\{0,1\}^{<\mathbb{N}}$ be the set of finite sequences of zeros and ones, and $\{0,1\}^{\mathbb{N}}$ the Cantor space of infinite such sequences. For $t \in \{0,1\}^{<\mathbb{N}}$, write $|t|$ for the length of t and $t \sqsubset \sigma$ when $\sigma \in \{0,1\}^{\mathbb{N}}$ extends t. The function $v_t\colon \{0,1\}^{\mathbb{N}} \to \{0,1\}$ is 1 on the clopen set $V_t = \{\sigma \in \{0,1\}^{\mathbb{N}} : t \sqsubset \sigma\}$ and 0 elsewhere. The points $t^\frown 0^\infty$ and $t^\frown 1^\infty$ extend t by zeros and by ones, respectively. Let $\prec$ be the lexicographic order on $\{0,1\}^{\mathbb{N}}$, with $\preceq$ the reflexive version. For $\sigma \in \{0,1\}^{\mathbb{N}}$, define $x^+_\sigma, x^-_\sigma\colon \{0,1\}^{\mathbb{N}} \to \{0,1\}$ as the characteristic functions of $\{\zeta : \sigma \preceq \zeta\}$ and $\{\zeta : \sigma \prec \zeta\}$, respectively. For topological spaces X_1, X_2, with disjoint sum $X_1 \oplus X_2$, and real functions $f\colon X_1 \to \mathbb{R}$ and $g\colon X_2 \to \mathbb{R}$, let $(f,g)\colon X_1 \oplus X_2 \to \mathbb{R}$ take the value $f(x)$ at the copy of $x \in X_1$ and $g(x)$ at the copy of $x \in X_2$. Since a point of $\{0,1\}^{\mathbb{N}}$ is a function from $\mathbb{N}$ to $\{0,1\}$, the Cantor space is a subset of $\mathbb{R}^{\mathbb{N}}$. The compacta below keep the names under which Argyros, Dodos and Kanellopoulos introduced them, such as $A(2^{\mathbb{N}})$, $\hat{D}(2^{\mathbb{N}})$ and $A(2^{<\mathbb{N}})$, in which $2^{\mathbb{N}}$ stands for the Cantor space and $2^{<\mathbb{N}}$ for the set of finite sequences.

The following compacta appear in the classification:

- the *one-point compactification* $A(2^{\mathbb{N}})$ of a discrete set of size continuum, realized as $\{\delta_\sigma : \sigma \in \{0,1\}^{\mathbb{N}}\} \cup \{0\}$, where δ_σ is the characteristic function of $\{\sigma\}$, and the separable extension $\hat{A}(2^{\mathbb{N}})$ of $A(2^{\mathbb{N}})$, the pointwise closure of $\{v_t : t \in \{0,1\}^{<\mathbb{N}}\}$, which equals $\{v_t : t \in \{0,1\}^{<\mathbb{N}}\} \cup A(2^{\mathbb{N}})$;
- the *split Cantor set* $S(2^{\mathbb{N}}) = \{x^+_\sigma, x^-_\sigma : \sigma \in \{0,1\}^{\mathbb{N}}\}$, the *split interval* $S(I)$, that is, $I \times \{0,1\}$ for the unit interval $I = [0,1]$, ordered lexicographically, with the order topology, and the extended split Cantor sets $\hat{S}_+(2^{\mathbb{N}})$ and $\hat{S}_-(2^{\mathbb{N}})$ defined below; the space $\hat{S}_+(2^{\mathbb{N}})$ is homeomorphic to a closed subspace of $S(I)$;
- the *Alexandroff duplicate* $D(2^{\mathbb{N}})$ of the Cantor set and the separable extension $\hat{D}(2^{\mathbb{N}})$ of $D(2^{\mathbb{N}})$, and the extended duplicate $\hat{D}(S(2^{\mathbb{N}}))$ of the split Cantor set, the last two being defined below; the duplicate $D(2^{\mathbb{N}})$ is the set $\{0,1\}^{\mathbb{N}} \times \{0,1\}$ in which the points $(\sigma,1)$ are isolated and the basic neighborhoods of a point $(\sigma,0)$ are the sets $(U \times \{0,1\}) \setminus \{(\sigma,1)\}$, for open neighborhoods U of σ in $\{0,1\}^{\mathbb{N}}$.

Among these, the space $S(2^{\mathbb{N}})$ is hereditarily separable but not metrizable, the space $D(2^{\mathbb{N}})$ is first countable but not separable, and the space $A(2^{\mathbb{N}})$ is not first countable (see [**ADK08**]).

The heptachotomy. A *dyadic subtree* of $\{0,1\}^{<\mathbb{N}}$ is a subset of $\{0,1\}^{<\mathbb{N}}$ that, ordered by extension, is isomorphic to $\{0,1\}^{<\mathbb{N}}$; we write the subtree as $(s_t)_{t\in\{0,1\}^{<\mathbb{N}}}$, using the canonical enumeration, the unique such isomorphism $t \mapsto s_t$ that also preserves the lexicographic order (see [ADK08, Section 2]). A dyadic subtree is *regular* if each level of the subtree is contained in a level of $\{0,1\}^{<\mathbb{N}}$. Fix a regular dyadic subtree $\Theta = (s_t)_{t\in\{0,1\}^{<\mathbb{N}}}$ such that the sequences $s^\frown 0^\infty$, for $s \in \Theta$, are pairwise distinct, and so are the sequences $s^\frown 1^\infty$; a family indexed by the nodes s of Θ is regarded as indexed by $\{0,1\}^{<\mathbb{N}}$ through the canonical enumeration $t \mapsto s_t$.

Definition P.14. Let X be a Polish space.

(1) For topological spaces X_1, X_2 and a countable index set J, let $f_j\colon X_1 \to \mathbb{R}$ and $g_j\colon X_2 \to \mathbb{R}$ be functions for $j \in J$, with each family pointwise bounded. The families $\{f_j : j \in J\}$ and $\{g_j : j \in J\}$ are *equivalent* if the assignment $f_j \mapsto g_j$ is well-defined and extends to a homeomorphism from the closure of $\{f_j : j \in J\}$ in $\mathbb{R}^{X_1}$ onto the closure of $\{g_j : j \in J\}$ in $\mathbb{R}^{X_2}$, both with the product topology.

(2) A family of pairwise distinct functions $\{f_t : t \in \{0,1\}^{<\mathbb{N}}\}$ that is relatively compact in $B_1(X)$ is *minimal* if the family is equivalent to $\{f_{s_t} : t \in \{0,1\}^{<\mathbb{N}}\}$ for every dyadic subtree (s_t) of $\{0,1\}^{<\mathbb{N}}$.

The seven minimal families are the following, and the pointwise closure of each of them is a separable Rosenthal compactum that contains the family as a discrete set (see [ADK08]):

(1) $D_1 = \{\frac{1}{|t|+1} v_t : t \in \{0,1\}^{<\mathbb{N}}\}$, whose closure $A(2^{<\mathbb{N}})$ is the one-point compactification of a countable discrete set;
(2) $D_2 = \{s^\frown 0^\infty : s \in \Theta\}$, a family of points of $\{0,1\}^{\mathbb{N}} \subseteq \mathbb{R}^{\mathbb{N}}$, whose closure is denoted $2^{\leq\mathbb{N}}$;
(3) $D_3 = \{x^+_{s^\frown 0^\infty} : s \in \Theta\}$, whose closure is $\hat{S}_+(2^{\mathbb{N}})$;
(4) $D_4 = \{x^-_{s^\frown 1^\infty} : s \in \Theta\}$, whose closure is $\hat{S}_-(2^{\mathbb{N}})$;
(5) $D_5 = \{v_t : t \in \{0,1\}^{<\mathbb{N}}\}$, whose closure is $\hat{A}(2^{\mathbb{N}})$;
(6) $D_6 = \{(v_t, t^\frown 0^\infty) : t \in \{0,1\}^{<\mathbb{N}}\}$, whose closure is $\hat{D}(2^{\mathbb{N}})$;
(7) $D_7 = \{(v_s, x^+_{s^\frown 0^\infty}) : s \in \Theta\}$, whose closure is $\hat{D}(S(2^{\mathbb{N}}))$.

The closures in (1) and (2) are metrizable, those in (3) and (4) are hereditarily separable but not metrizable, those in (6) and (7) are first countable but not hereditarily separable, and the closure in (5) is not first countable (see [ADK08]).

Fact P.15 (Heptachotomy; see [ADK08, Theorem 2]).

(a) *Up to equivalence, there are exactly seven minimal families, namely $D_1, \dots, D_7$, and they are pairwise non-equivalent.*

(b) *Let X be a Polish space and $\{f_t : t \in \{0,1\}^{<\mathbb{N}}\}$ a family of pairwise distinct functions that is relatively compact in $B_1(X)$. Then there is a regular dyadic subtree (s_t) of $\{0,1\}^{<\mathbb{N}}$ such that $\{f_{s_t} : t \in \{0,1\}^{<\mathbb{N}}\}$ is equivalent to one of $D_1, \dots, D_7$.*

Todorčević's trichotomy. Let X be a Polish space, $K \subseteq B_1(X)$ a separable Rosenthal compactum and $\{f_n : n \in \mathbb{N}\}$ a countable dense subset of K. An injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ is *increasing* if $\iota(s) < \iota(t)$ whenever s precedes t in the order of $\{0,1\}^{<\mathbb{N}}$ by length and, within each length, by the lexicographic order. A minimal family D_i *canonically embeds* into K (with respect to $\{f_n\}$) if there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that $\{f_{\iota(t)} : t \in \{0,1\}^{<\mathbb{N}}\}$ is equivalent to D_i. In that case the closure of D_i is homeomorphic to a closed subspace of K.

Fact P.16 (Todorčević's trichotomy in canonical form; see [Tod99] and [ADK08, Theorem 3]). *Let X be a Polish space, $K \subseteq B_1(X)$ a separable Rosenthal compactum and $\{f_n : n \in \mathbb{N}\}$ a countable dense subset of K.*

(a) *If K consists of bounded functions, is metrizable and is not separable in the supremum norm, then the family D_2 canonically embeds into K.*
(b) *If K is hereditarily separable but not metrizable, then the family D_3 or the family D_4 canonically embeds into K.*
(c) *If K is first countable but not hereditarily separable, then the family D_6 or the family D_7 canonically embeds into K.*
(d) *If K is not first countable, then the family D_5 canonically embeds into K.*

In particular, a non-metrizable separable Rosenthal compactum contains a copy of $\hat{S}_+(2^{\mathbb{N}})$ or $\hat{S}_-(2^{\mathbb{N}})$, of $\hat{D}(2^{\mathbb{N}})$ or $\hat{D}(S(2^{\mathbb{N}}))$, or of $\hat{A}(2^{\mathbb{N}})$, according as the compactum is hereditarily separable, first countable but not hereditarily separable, or not first countable. If K is metrizable and infinite, then the family D_1 canonically embeds into K, as recorded in Figure 4. Indeed, an infinite compact metrizable space has a non-isolated point g. Since the enumerated family is dense, we can choose indices $n_0 < n_1 < \cdots$ such that the functions f_{n_k} are pairwise distinct, differ from g and converge to g. Enumerate $\{0,1\}^{<\mathbb{N}}$ as $t_0, t_1, \ldots$ by length and, within each length, lexicographically, and let $\iota(t_k) = n_k$, which defines an increasing injection. The closure of $\{f_{\iota(t)} : t \in \{0,1\}^{<\mathbb{N}}\}$ and the closure $D_1 \cup \{0\}$ of D_1 each consist of a sequence of pairwise distinct isolated points together with the limit of the sequence, and every bijection between the isolated points of two such spaces extends, by sending limit to limit, to a homeomorphism. Hence the family $\{f_{\iota(t)} : t \in \{0,1\}^{<\mathbb{N}}\}$ is equivalent to D_1.

Levels of NIP.. We now transfer the classification to NIP families, as in our work on deep computations [**DIMW+26**, Section 4.2].

Definition P.17. Let Y be a compact metrizable space and $\mathcal{R} \subseteq C_p(Y)$ a pointwise bounded NIP family whose closure $\overline{\mathcal{R}}$ is separable. The family $\mathcal{R}$ is *NIP*$^{\mathrm{fc}}$, *NIP*$^{\mathrm{hs}}$ or *NIP*$^{\mathrm{met}}$ if $\overline{\mathcal{R}}$ is, respectively, first countable, hereditarily separable or metrizable.

By the implications among the three invariants, we have NIP$^{\mathrm{met}}$ $\Rightarrow$NIP$^{\mathrm{hs}}$ $\Rightarrow$NIP$^{\mathrm{fc}}$ $\Rightarrow$NIP.

The top level has a concrete meaning. A set $Z \subseteq Y$ *separates* a set K of functions on Y if any two distinct members of K differ at some point of Z.

Proposition P.18 (see [Eng89]). *Let Y be a set and $K \subseteq \mathbb{R}^Y$ a set that is compact in the topology of pointwise convergence. Then the space K is metrizable if and only if some countable set $Z \subseteq Y$ separates K.*

Thus an NIP family with separable closure is NIP$^{\mathrm{met}}$ exactly when countably many points of Y distinguish all functions in the closure of the family. When the functions are responses and the points are probes, this says that countably many probes distinguish all idealized responses. The other two levels have readings of the same kind, for which we need two notions. Let K be a set of real functions on a set Y. A member f of K is *countably determined* if there is a countable set $Z_f \subseteq Y$ such that f is the only member of K that agrees with f on Z_f. A set $E \subseteq K$ is *finitely separated* if, for each $f \in E$, there are a finite set $F_f \subseteq Y$ and a real number $\varepsilon_f > 0$ such that every $g \in E \setminus \{f\}$ satisfies $|g(y) - f(y)| \geq \varepsilon_f$ for some $y \in F_f$.

Proposition P.19. *Let Y be a compact metrizable space and $\mathcal{R} \subseteq C_p(Y)$ a pointwise bounded NIP family whose closure $K = \overline{\mathcal{R}}$ is separable.*

(a) *The family is NIP*$^{\mathrm{met}}$ *if and only if some countable set $Z \subseteq Y$ separates K.*
(b) *The family is NIP*$^{\mathrm{fc}}$ *if and only if every member of K is countably determined.*
(c) *The family is NIP*$^{\mathrm{hs}}$ *if and only if no subset of K of cardinality $2^{\aleph_0}$ is finitely separated.*

Proof. Part (a) is Proposition P.18. Part (b) holds because, in a compact Hausdorff space, a point has a countable neighborhood base exactly when the singleton of the point is a countable intersection of open sets. Part (c) is Pol's theorem, by which a separable Rosenthal compactum is hereditarily separable exactly when the compactum has no discrete subspace of cardinality $2^{\aleph_0}$ (see [**ADK08**] and [**GM22**, Theorem 0.4]); a set is finitely separated exactly when it is a discrete subspace. □

The three levels thus measure how many points of Y are needed to determine a limit in the closure K: one countable set for all limits ($\mathrm{NIP}^{\mathrm{met}}$), a countable set for each limit ($\mathrm{NIP}^{\mathrm{fc}}$), or no family of size continuum in which each member can be isolated by finitely many points at a positive tolerance ($\mathrm{NIP}^{\mathrm{hs}}$). In the last condition, both the finite set and the tolerance may depend on the member being isolated. When limits are taken as types for a fixed formula and points as parameters, the levels describe how instances of the formula determine a type. This partly answers a question, raised in our work on deep computations [**DIMW$^+$26**], that asks for a non-topological characterization of the levels; a finitary characterization, in the spirit of the fat-shattering dimension, remains open. The fat-shattering dimension at a given gap is the supremum of the sizes of the finite sets of points at which the functions realize every pattern of high and low values, around thresholds that may vary from point to point.

Theorem P.20 ([DIMW$^+$26, Section 4.2]). *Let Y be a compact metrizable space and $\{r_n : n \in \mathbb{N}\} \subseteq C_p(Y)$ a pointwise bounded NIP family. Then exactly one of the following holds:*

(i) *the family is $\mathrm{NIP}^{\mathrm{met}}$;*
(ii) *the family is $\mathrm{NIP}^{\mathrm{hs}}$ but not $\mathrm{NIP}^{\mathrm{met}}$, and there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that $\{r_{\iota(t)} : t \in \{0,1\}^{<\mathbb{N}}\}$ is equivalent to D_3 or to D_4;*
(iii) *the family is $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$, and there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that $\{r_{\iota(t)} : t \in \{0,1\}^{<\mathbb{N}}\}$ is equivalent to D_6 or to D_7;*
(iv) *the family is not $\mathrm{NIP}^{\mathrm{fc}}$, and there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that $\{r_{\iota(t)} : t \in \{0,1\}^{<\mathbb{N}}\}$ is equivalent to D_5.*

The theorem combines Theorem P.11, Lemma P.7 and Fact P.16(b)–(d), applied to the closure of the family and to the family as a countable dense subset.

Thus a non-metrizable NIP closure is always explained by a subfamily of the original functions, arranged along the dyadic tree and equivalent to one of five minimal families. Applied to realized types, the complete descriptions of elements of the structure in terms of the parameters, the classification of Theorem P.20 yields the classification of NIP formulas given by Dueñez, Iovino, Matos-Wiederhold, Salvetti and Tall [**DIMW$^+$26**, Section 4.2]. There the realized types are viewed as continuous functions on a compact metrizable space of parameters, and this proviso matters: the space of types of a formula over a countable set of parameters, with no topology on the parameters, is always metrizable, since the countable set of parameters separates the types (see Proposition P.18).

The idealized responses of an observable form a compact set of functions. To apply the classification to a bipartite observable, we retain from Bob's operator algebra the effective observables produced by Alice's preparations. The states of the smallest unital C*-algebra containing them form a compact metrizable probe space, on which each actual preparation has a continuous response. We identify the pointwise closure of these responses with the responses obtained from all states of Alice's operator algebra, with Bob's system idealized first. The compactum to be classified thus records preparations through their responses to the fixed observable.

Observables and responses. We use the notation of Variatio 1 and of section P.1: the space $\mathfrak{T}(H)$ of trace-class operators on a Hilbert space H, the set $\mathfrak{D}(H)$ of density operators, the expectation φ_C of an observable C on $H_1 \otimes H_2$, the response $r_\rho\colon \mathfrak{D}(H_2) \to [-\|C\|, \|C\|]$ of a density operator ρ on actual probes, given by $r_\rho(\sigma) = \varphi_C(\rho, \sigma)$, and the slice map $T_C\colon \mathfrak{T}(H_1) \to B(H_2)$, of norm at most $\|C\|$, which sends a density operator ρ to the effective observable $T_C(\rho)$. A *C*-algebra* is a norm-closed algebra of bounded operators closed under adjoints; we use only unital ones, which contain the identity. A *state* of a unital C*-algebra $\mathfrak{A}$ is a linear functional ω on $\mathfrak{A}$ that is positive, in the sense that $\omega(a^*a) \geq 0$ for all $a \in \mathfrak{A}$, and satisfies $\omega(I) = 1$. The set $\mathcal{S}(\mathfrak{A})$ of states, with the weak* topology, is compact, and metrizable when $\mathfrak{A}$ is norm-separable. By Fact P.1, the actual states are weak* dense in $\mathcal{S}(B(H))$, and every functional in $B(H)^*$ is a complex linear combination of four states.

The space of probes. The responses of Alice's preparations involve Bob's system only through the effective observables $T_C(\rho)$. The natural probes are therefore the states of the smallest unital C*-algebra containing them.

Definition P.21. Let C be an observable on $H_1 \otimes H_2$. The *probe algebra* $\mathfrak{A}_C \subseteq B(H_2)$ is the unital C*-algebra generated by $T_C(\mathfrak{T}(H_1))$, and the *probe space* is $Y_C = \mathcal{S}(\mathfrak{A}_C)$. The *record* of an actual probe $\sigma \in \mathfrak{D}(H_2)$ is the functional $\mathfrak{p}_\sigma\colon \mathfrak{A}_C \to \mathbb{C}$ given by $\mathfrak{p}_\sigma(a) = \operatorname{Tr}(\sigma a)$. For $\rho \in \mathfrak{D}(H_1)$, the *completed response* of ρ is the function $\mathfrak{r}_\rho\colon Y_C \to [-\|C\|, \|C\|]$ given by $\mathfrak{r}_\rho(\omega_2) = \omega_2(T_C(\rho))$, and the *response family* is $\mathcal{R}_C = \{\mathfrak{r}_\rho : \rho \in \mathfrak{D}(H_1)\}$.

The record $\mathfrak{p}_\sigma$ is a state of $\mathfrak{A}_C$, and hence a point of Y_C, rather than a density operator; different density operators may have the same record. In a completed response the preparation ρ remains actual: only the domain of probes is completed.

Lemma P.22. *Let C be an observable on $H_1 \otimes H_2$.*

(a) *The probe space Y_C is compact and metrizable.*
(b) *For every $\rho \in \mathfrak{D}(H_1)$, the completed response $\mathfrak{r}_\rho$ is continuous on Y_C and bounded by $\|C\|$, so the response family $\mathcal{R}_C$ is a pointwise bounded subset of $C_p(Y_C)$.*
(c) *The records $\mathfrak{p}_\sigma$ of actual probes $\sigma \in \mathfrak{D}(H_2)$ form a dense subset of Y_C, and for $\rho \in \mathfrak{D}(H_1)$ and $\sigma \in \mathfrak{D}(H_2)$ we have $\mathfrak{r}_\rho(\mathfrak{p}_\sigma) = r_\rho(\sigma) = \varphi_C(\rho, \sigma)$.*

Proof. (a) Because H_1 is separable, so is $\mathfrak{T}(H_1)$; the image of $\mathfrak{T}(H_1)$ under the norm-continuous slice map is therefore separable, and so is the C*-algebra $\mathfrak{A}_C$ that this image generates. The compact space Y_C is thus metrizable.

(b) The response $\mathfrak{r}_\rho$ is evaluation at the element $T_C(\rho)$ of $\mathfrak{A}_C$, and evaluation at a fixed element is weak* continuous. Since $\rho \geq 0$ and C is self-adjoint, the operator $T_C(\rho)$ is self-adjoint, whence $\mathfrak{r}_\rho$ is real-valued; moreover, $|\mathfrak{r}_\rho(\omega_2)| \leq \|T_C(\rho)\| \leq \|C\|$.

(c) By the Hahn–Banach theorem, every state of $\mathfrak{A}_C$ extends to a state of $B(H_2)$ (see [**KR83**, Theorem 4.3.13]); restriction to $\mathfrak{A}_C$ is thus a weak*-continuous map of the state space of $B(H_2)$ onto Y_C. The states given by density operators are weak* dense in the state space of $B(H_2)$ (see section P.2), and their restrictions are therefore dense in Y_C. The final identity restates the characterization of the effective observable. □

By Lemma P.22(c) and Lemma P.12, Bob's actual probes alone decide whether $\mathcal{R}_C$ has the independence property. The family $\mathcal{R}_C$ has the independence property exactly when there are actual states $\rho_1, \rho_2, \ldots$ of Alice's system and reals $a < b$ with the following property: for all disjoint finite $E, F \subseteq \mathbb{N}$, some actual state σ of Bob's system satisfies $\varphi_C(\rho_n, \sigma) \leq a$ for $n \in E$ and $\varphi_C(\rho_n, \sigma) \geq b$ for $n \in F$. A single probe then lets Bob read any finite pattern off Alice's sequence.

Definition P.23. An observable C on $H_1 \otimes H_2$ is *NIP for Alice's responses* if the response family $\mathcal{R}_C$ is NIP.

Idealized responses. To describe the pointwise closure of $\mathcal{R}_C$, we extend the slice map from trace-class operators to functionals on $B(H_1)$. For a bounded operator $T\colon V \to W$ between Banach spaces, the transpose of T is $T'\colon W^* \to V^*$, given by $T'(\lambda) = \lambda \circ T$, and the bitranspose is $T''\colon V^{**} \to W^{**}$. Under the canonical embeddings of each Banach space into the bidual, the bitranspose extends T and is weak*-to-weak* continuous. We regard the slice map as a map into $\mathfrak{A}_C$, so that the relevant maps are $T'_C\colon \mathfrak{A}_C^* \to B(H_1)$ and $T''_C\colon B(H_1)^* \to \mathfrak{A}_C^{**}$, where $\mathcal{T}(H_1)^*$ is identified with $B(H_1)$. The canonical image of a density operator ρ in $B(H_1)^*$ is the functional $a \mapsto \operatorname{Tr}(\rho a)$; the same domain also contains every idealized state of $B(H_1)$.

For a fixed observable C on $H_1 \otimes H_2$ and a state $\omega_1 \in \mathcal{S}(B(H_1))$, define $\mathfrak{r}_{\omega_1}\colon Y_C \to [-\|C\|, \|C\|]$ by

$$\mathfrak{r}_{\omega_1}(\omega_2) = \langle T''_C(\omega_1), \omega_2 \rangle = \omega_1(T'_C\omega_2).$$

For each $\omega_2 \in Y_C$, the operator $T'_C\omega_2$ is self-adjoint with norm at most $\|C\|$, so the displayed value is real and satisfies the stated bound. The response depends affinely on each state separately, and when ω_1 is the state represented by a density operator ρ the response agrees with $\mathfrak{r}_\rho$ of Definition P.21. We recall from Variatio 1 and (1) the two products of $\omega_1 \in \mathcal{S}(B(H_1))$ and $\widetilde{\omega}_2 \in \mathcal{S}(B(H_2))$, evaluated at C. Let $S_C\colon \mathcal{T}(H_2) \to B(H_1)$ be the slice map on Bob's side, $S_C(y) = \operatorname{Tr}_2((I \otimes y)C)$, so that $\varphi_C(\rho, \sigma) = \operatorname{Tr}(\rho\, S_C(\sigma))$. The products are $(\omega_1 \otimes_{\mathsf{A}} \widetilde{\omega}_2)(C) = \widetilde{\omega}_2(Z)$, where $Z \in B(H_2)$ is given by $\operatorname{Tr}(yZ) = \omega_1(S_C y)$ for $y \in \mathcal{T}(H_2)$, and $(\omega_1 \otimes_{\mathsf{B}} \widetilde{\omega}_2)(C) = \omega_1(W)$, where $W \in B(H_1)$ is given by $\operatorname{Tr}(xW) = \widetilde{\omega}_2(T_C x)$ for $x \in \mathcal{T}(H_1)$. Taking first the limit of the state inside the slice map, and then the limit of the other state, shows that for nets $(\rho_i)_{i\in I}$ and $(\sigma_j)_{j\in J}$ of actual states, indexed by directed sets I and J and converging weak* to ω_1 and $\widetilde{\omega}_2$, respectively, the products satisfy

$$(\omega_1 \otimes_{\mathsf{A}} \widetilde{\omega}_2)(C) = \lim_j \lim_i \varphi_C(\rho_i, \sigma_j), \qquad (\omega_1 \otimes_{\mathsf{B}} \widetilde{\omega}_2)(C) = \lim_i \lim_j \varphi_C(\rho_i, \sigma_j).$$

The subscript names the system idealized first, that is, the one whose limit is taken inside.

Lemma P.24. *Let C be an observable on $H_1 \otimes H_2$. Then*

$$\overline{\mathcal{R}_C} = \{\mathfrak{r}_{\omega_1} : \omega_1 \in \mathcal{S}(B(H_1))\}.$$

Moreover, if ω_1 is a state of $B(H_1)$, $\omega_2 \in Y_C$ and $\widetilde{\omega}_2$ is a state of $B(H_2)$ extending ω_2, then $\mathfrak{r}_{\omega_1}(\omega_2) = (\omega_1 \otimes_{\mathsf{B}} \widetilde{\omega}_2)(C)$.

Proof. For each probe $\omega_2 \in Y_C$, the coordinate $\omega_1 \mapsto \mathfrak{r}_{\omega_1}(\omega_2) = \omega_1(T'_C\omega_2)$ is weak* continuous on $\mathcal{S}(B(H_1))$. Hence the response map $\omega_1 \mapsto \mathfrak{r}_{\omega_1}$ is continuous into $\mathbb{R}^{Y_C}$ with the pointwise topology, and the image K of the map is compact. Since the bitranspose extends the slice map, the image contains $\mathcal{R}_C$ and therefore contains the pointwise closure of $\mathcal{R}_C$. Conversely, actual states are weak* dense in $\mathcal{S}(B(H_1))$ (see section P.2), so continuity gives $K \subseteq \overline{\mathcal{R}_C}$. For the final identity, if $\widetilde{\omega}_2 \in \mathcal{S}(B(H_2))$ restricts to $\omega_2 \in Y_C$, then

$$\mathfrak{r}_{\omega_1}(\omega_2) = \omega_1(T'_C\omega_2) = (\omega_1 \otimes_{\mathsf{B}} \widetilde{\omega}_2)(C),$$

by the definition of the product: since T_C takes values in $\mathfrak{A}_C$, the functional $x \mapsto \widetilde{\omega}_2(T_C x)$ on $\mathcal{T}(H_1)$ equals $x \mapsto \omega_2(T_C x)$, which is $T'_C\omega_2$. Thus the value is independent of the extension $\widetilde{\omega}_2$. □

We call the members of $\overline{\mathcal{R}_C}$ the *idealized responses*; they include the completed responses of actual preparations. By Lemma P.24, the response map $\omega_1 \mapsto \mathfrak{r}_{\omega_1}$ sends $\mathcal{S}(B(H_1))$ continuously onto this compactum; different states may have the same response. A probe in Y_C,

the restriction of a state of $B(H_2)$, need not be represented by a density operator, and an idealized response evaluates the probe with Bob's system idealized first, whichever extension of the probe is chosen. On actual probes the two orders agree: for $\sigma \in \mathfrak{D}(H_2)$ the operator W above is $S_C(\sigma)$, and both products equal $\omega_1(S_C\sigma)$; thus, for every state ω_1 of $B(H_1)$ and every $\sigma \in \mathfrak{D}(H_2)$, we have $\mathfrak{r}_{\omega_1}(\mathfrak{p}_\sigma) = (\omega_1 \otimes_{\mathsf{B}} \sigma)(C) = (\omega_1 \otimes_{\mathsf{A}} \sigma)(C)$.

Stability, NIP and the independence property are continuity, membership in the first Baire class and a copy of $\beta\mathbb{N}$. Stability and NIP admit a common description in terms of the regularity of idealized responses. Recall from Variatio 2 that an observable is stable when, for all sequences of actual preparations of the two systems, the two iterated limits of the expectation values, one taken first over Alice's sequence and the other first over Bob's, agree whenever both exist. In this sense the expectation value is independent of the order in which the two systems are idealized. Here an idealized response is the expectation of a preparation of Alice's system, actual or idealized, as a function of Bob's probe, with Bob's system idealized first. In Theorem 11.1, we characterize stability by continuity of every idealized response and NIP by membership of every idealized response in the first Baire class; when NIP fails, a sequence of actual responses has a closure homeomorphic to the space $\beta\mathbb{N}$ of ultrafilters on the natural numbers. Within NIP, the finer classification concerns the topology of the compactum formed by these functions, and the affine dependence on preparations imposes a further restriction through convexity.

To state the result, we recall the notation of section P.2: an observable C is a bounded self-adjoint operator on the tensor product $H_1 \otimes H_2$ of the Hilbert spaces of Alice's and Bob's systems, the probe space Y_C is the compact metrizable state space of the probe algebra (see Definition P.21 and Lemma P.22), and the response family $\mathcal{R}_C$ consists of the responses of Alice's actual preparations; the closure $\overline{\mathcal{R}_C}$ of this family is the set of idealized responses (see Lemma P.24).

For stability we need one more fact from operator theory, Kadison's function representation: on the state space of a unital C*-algebra $\mathfrak{A}$, the continuous affine real functions are exactly the evaluations $\omega_2 \mapsto \omega_2(z)$ at self-adjoint elements $z \in \mathfrak{A}$ (see [**Kad51**]).

Proof of Theorem 11.1. Part (a) rests on the characterization of stability by weak compactness, and parts (b) and (c) rest on the dichotomy of Theorem P.11. Recall that a bounded operator between Banach spaces is *weakly compact* if it maps bounded sets to relatively weakly compact sets, and that, by Gantmacher's theorem [**Gan40**], this holds exactly when the bitranspose of the operator takes values in the original target space. The observable C is stable exactly when the slice map $T_C\colon \mathfrak{T}(H_1) \to B(H_2)$ is weakly compact (see Theorem 5.1 and its proof). The equivalence also follows directly from Grothendieck's double limit criterion [**Gro52**]: a bounded set A in the dual of a Banach space Y is relatively weakly compact exactly when $\lim_k \lim_l a_k(y_l) = \lim_l \lim_k a_k(y_l)$ for all sequences $(a_k)_{k\in\mathbb{N}}$ in A and $(y_l)_{l\in\mathbb{N}}$ in the unit ball of Y for which both iterated limits exist. We apply the criterion with $Y = \mathfrak{T}(H_2)$ and A the image under T_C of the unit ball of $\mathfrak{T}(H_1)$, so that $a_k(y_l) = \operatorname{Tr}(y_l\, T_C x_k)$. Every operator in the unit ball of either trace class is a combination $c_1\tau_1 - c_2\tau_2 + ic_3\tau_3 - ic_4\tau_4$ of density operators τ_j with $0 \le c_j \le 1$; a diagonal passage to subsequences makes the coefficients converge and makes both iterated limits exist for each of the finitely many pairs of component sequences of density operators. By stability, the two iterated limits of each pair agree, and their linear combination gives the equality required by the criterion. Conversely, the criterion restricted to density operators is the stability of C.

For part (a), suppose first that C is stable. Then the slice map is weakly compact as a map into $B(H_2)$, and therefore also as a map into the closed subspace $\mathfrak{A}_C$, whose weak topology

is the restriction of that of $B(H_2)$. The bitranspose T_C'' then maps $B(H_1)^*$ into $\mathfrak{A}_C$, whence every idealized response $\mathfrak{r}_{\omega_1}$ is evaluation at an element of $\mathfrak{A}_C$ and is continuous on Y_C.

Conversely, suppose that every idealized response is continuous, and let $\omega_1 \in \mathcal{S}(B(H_1))$ be given. The response $\mathfrak{r}_{\omega_1}$ is real and affine on Y_C by the defining formula, so Kadison's function representation gives a self-adjoint $z \in \mathfrak{A}_C$ with $\mathfrak{r}_{\omega_1}(\omega_2) = \omega_2(z)$ for every $\omega_2 \in Y_C$. Thus the elements $T_C''(\omega_1)$ and z of $\mathfrak{A}_C^{**}$ agree on all states of $\mathfrak{A}_C$; since these states span $\mathfrak{A}_C^*$, the elements are equal. Every functional in $B(H_1)^*$ is a complex linear combination of states (see section P.2), so linearity implies $T_C''(B(H_1)^*) \subseteq \mathfrak{A}_C$. Gantmacher's theorem therefore makes $T_C \colon \mathcal{T}(H_1) \to \mathfrak{A}_C$ weakly compact. Composing with the bounded inclusion $\mathfrak{A}_C \hookrightarrow B(H_2)$ preserves weak compactness, and the criterion above implies that C is stable.

By Lemma P.22, the dichotomy of Theorem P.11 applies to the family $\mathcal{R}_C \subseteq C_p(Y_C)$, and part (b) follows from the dichotomy together with Lemma P.7. For part (c), let actual states $\rho_n \in \mathfrak{D}(H_1)$, for $n \in \mathbb{N}$, and reals $a < b$ witness the independence property of $\mathcal{R}_C$. By compactness of Y_C (see the proof of Theorem P.11), for every $I \subseteq \mathbb{N}$ some probe ω_2 satisfies $\mathfrak{r}_{\rho_n}(\omega_2) \le a$ for $n \in I$ and $\mathfrak{r}_{\rho_n}(\omega_2) \ge b$ for $n \notin I$. No subsequence of $(\mathfrak{r}_{\rho_n})$ can therefore converge pointwise, and Fact P.9 yields a subsequence whose closure is homeomorphic to $\beta\mathbb{N}$; the states of this subsequence form the required sequence. □

In terms of preparations and probes, Theorem 11.1 describes how the responses of Alice's preparations depend on Bob's probe at each level:

- for a stable observable, the response of each idealized preparation of Alice's system depends continuously on Bob's probe, idealized probes included;
- for an observable that is NIP but unstable, such as the comparison of photon numbers (see section P.2), these responses can be discontinuous, but each is a pointwise limit of a sequence of continuous functions, and together they form a Rosenthal compactum;
- for an observable with the independence property, a sequence of Alice's actual preparations lets Bob read every finite pattern with a single probe, and every such sequence has a subsequence whose closure is a copy of $\beta\mathbb{N}$ (see the proof of Theorem 11.1(c)).

Remark P.25. As stability has a uniform form (see Variatio 7), so has NIP. For a real number $\delta > 0$, a finite set of functions on a set Y is *δ-shattered* if there are reals $a < b$ with $b - a \ge \delta$ such that, for any two disjoint subsets E and F of the set, some point $y \in Y$ satisfies $r(y) \le a$ for $r \in E$ and $r(y) \ge b$ for $r \in F$. The notion of NIP in Definition P.8 forbids an infinite sequence all of whose finite patterns are realized with fixed reals $a < b$, which we call an *independent* sequence; *uniform NIP* bounds, for each $\delta > 0$, the size of the δ-shattered subsets of the family, and so implies NIP. For a uniformly bounded family, such as a family of responses, uniform NIP amounts, by the pigeonhole principle applied to the thresholds and by Assouad's duality argument (see Remark P.13), to the finiteness of the fat-shattering dimension at every gap.

Comparisons $[A < B]$, the projections that represent the event "Bob's value of B exceeds Alice's value of A" for quantities A of Alice's system and B of Bob's, are uniformly NIP, with bounds independent of the Hilbert spaces (see Corollary P.36), and they are unstable when both quantities have pure point spectra unbounded above, by the ladder argument of section P.2. By Theorem 11.1, their idealized responses form a Rosenthal compactum of Baire-1 functions, not all continuous.

As in Part I, the uniform notion is the one that survives ultrapowers. Recall that, for an ultrafilter $\mathcal{U}$ on $\mathbb{N}$, every bounded sequence $(x_k)_{k\in\mathbb{N}}$ of complex numbers has a unique ultralimit $\lim_{k,\mathcal{U}} x_k$, the complex number x such that $\{k : |x_k - x| < \varepsilon\} \in \mathcal{U}$ for every $\varepsilon > 0$. The extension of a family of functions on Y to the ultrapower by $\mathcal{U}$ consists of the functions

that send a sequence $(y_k)_{k\in\mathbb{N}}$ of points of Y to the ultralimit $\lim_{k,\mathcal{U}} r_k(y_k)$, one function for each sequence $(r_k)_{k\in\mathbb{N}}$ of members of the family; each value is a single ultralimit along the common index k, not the value of a limit function at a limit point. A uniformly bounded family is uniformly NIP if and only if the extension of the family to the ultrapower by every nonprincipal ultrafilter on $\mathbb{N}$ is NIP. Indeed, by Łoś's theorem a δ-shattered set in the ultrapower yields δ'-shattered sets of the same size in the family for every $\delta' < \delta$; conversely, if the sizes of the δ-shattered sets are unbounded, rounding the thresholds to a grid of mesh $\delta/3$ and the pigeonhole principle give fixed reals $a < b$ that shatter sets of every finite size, and $\aleph_1$-saturation of the ultrapower combines them into an infinite independent sequence.

Uniform NIP also has a statistical interpretation. After normalizing responses to $[0, 1]$, consider independently sampled probe–outcome pairs from an unknown joint distribution. In agnostic learning with absolute loss, one seeks a predictor whose mean absolute error is within a prescribed tolerance of the infimum over the response class, with a prescribed probability of success. For the response classes considered here, distribution-free sample bounds of this kind exist exactly when the fat-shattering dimension is finite at every positive gap [**BLW96**, **ABDCBH97**], that is, exactly when the class is uniformly NIP. Part I instead studies prediction along an adaptive sequence of probes, where uniform stability bounds the number of mistakes, whether the threshold is fixed or varies from probe to probe.

The finer classification of NIP observables. For NIP observables, the classification of section P.2 applies to the idealized responses. We first check that they form a separable compactum.

Lemma P.26. *Let C be an observable on $H_1 \otimes H_2$ and $(\rho_n)_{n\in\mathbb{N}}$ a sequence of density operators that is dense in $\mathfrak{D}(H_1)$ for the trace norm. Then the responses $\mathfrak{r}_{\rho_n}$ are dense in $\overline{\mathcal{R}_C}$.*

Proof. For density operators ρ and ρ', we have $\sup_{\omega_2\in Y_C} |\mathfrak{r}_\rho(\omega_2) - \mathfrak{r}_{\rho'}(\omega_2)| \le \|T_C(\rho - \rho')\| \le \|C\|\,\|\rho - \rho'\|_1$. Every response is therefore a uniform limit, and hence a pointwise limit, of a sequence of responses $\mathfrak{r}_{\rho_n}$; the closure of $\{\mathfrak{r}_{\rho_n}\}$ thus contains $\mathcal{R}_C$, and with it $\overline{\mathcal{R}_C}$. Sequences such as (ρ_n) exist, because $\mathfrak{T}(H_1)$ is separable. □

Definition P.27. An observable C that is NIP for Alice's responses is *NIP*$^{\mathrm{fc}}$, *NIP*$^{\mathrm{hs}}$ or *NIP*$^{\mathrm{met}}$ if the response family $\mathcal{R}_C$ is, respectively, NIP$^{\mathrm{fc}}$, NIP$^{\mathrm{hs}}$ or NIP$^{\mathrm{met}}$ (see Definition P.17).

Proposition P.28. *Every stable observable is NIP$^{\mathrm{met}}$: the idealized responses form a compact metrizable space.*

Proof. By Theorem 11.1(a), every idealized response is continuous on the compact metrizable space Y_C, hence Baire-1, and the observable C is therefore NIP for Alice's responses by Theorem 11.1(b); the set of idealized responses is separable by Lemma P.26. A countable dense set of probes separates continuous functions, and Proposition P.18 shows that the idealized responses form a metrizable space. □

Mixtures force metrizability. The idealized responses of an observable form a convex set, because responses depend affinely on Alice's state (see the proof of Corollary P.31). For separable convex Rosenthal compacta, hereditary separability implies metrizability (see Theorem P.29 below). This is a special case of a known result on locally connected Rosenthal compacta, due to Todorčević and Gruenhage (see Remark P.30). The mechanism is visible in the split interval $S(I)$ of section P.2, which is homeomorphic to the set of indicator functions $\mathbf{1}_{(a,1]}$ and $\mathbf{1}_{[a,1]}$ on $[0, 1]$, for $0 \le a \le 1$. For each such a, the two indicators form a split pair, two functions that differ only at a, and a convex set of functions that contains both of them contains their midpoint, which takes the value $\frac{1}{2}$ at a. These midpoints form an uncountable discrete set: the midpoint for a point $b \ne a$ takes the value 0 or 1 at a.

Theorem P.29 (Todorčević, Gruenhage [Tod99, Gru88]). *Let Y be a set, and let $K \subseteq \mathbb{R}^Y$ be a separable compact convex set which, with the topology of pointwise convergence, is a Rosenthal compactum. If K is hereditarily separable, then K is metrizable.*

For convex compacta there is a short proof, which uses the affine coordinates of K; the locally connected case is discussed in Remark P.30.

Proof. By a theorem of Todorčević [Tod99], a Rosenthal compactum without an uncountable discrete subspace admits a continuous map, at most two-to-one, onto a compact metric space. Since K is hereditarily separable, there is such a map $f\colon K \to M$, for a compact metric space (M, d). For each positive integer n, continuity and compactness give finitely many basic open sets of $\mathbb{R}^Y$, each determined by finitely many coordinates, that cover K and on whose traces on K the values of f have diameter less than $1/n$. Let $Z \subseteq Y$ be the countable set of the coordinates used, for all n. If two points of K agree on Z, then each of these basic sets that contains one of them contains the other, so the distance between their images is less than $1/n$ for every n, and their images coincide. Hence each fiber of the restriction map $\pi\colon K \to \mathbb{R}^Z$ lies in a fiber of f, which has at most two points. A fiber of π is convex, and a convex set with two distinct points contains the segment between them; the fibers of π are therefore single points, and π is a continuous injection of the compact space K into the metrizable space $\mathbb{R}^Z$, hence a homeomorphism onto its image. □

Remark P.30. The same midpoints decide a claim about the split interval $S(I) = [0,1] \times \{0,1\}$, ordered lexicographically, with the order topology. In a remark following his Proposition 10 [God80], Godefroy states that the space $P(S(I))$ of Borel probability measures on $S(I)$, with the weak* topology, is hereditarily separable. This is not the case: for $0 < a < 1$, let $a^- = (a, 0)$ and $a^+ = (a, 1)$, and let m_a be the midpoint $\frac{1}{2}(\delta_{a^-} + \delta_{a^+})$ of the point masses at a^- and a^+. The set $\{s \in S(I) : s \leq a^-\}$ is clopen, so the indicator f_a of this set is continuous, and $m_a(f_a) = \frac{1}{2}$, while $m_b(f_a) \in \{0, 1\}$ for $b \neq a$. Hence the set $\{m_a : 0 < a < 1\}$ is an uncountable discrete subspace of $P(S(I))$, as in the example before Theorem P.29. Todorčević has pointed out to us that Theorem P.29 is known, and that the theorem holds in the larger class of locally connected Rosenthal compacta, which contains the convex ones: the theorem follows from his Theorems 5 and 7 [Tod99] and from an argument of Gruenhage [Gru88, Lemma 2.4]. The examples of section P.2 realize the other three of the four classes, the metrizable compacta and the two classes of Todorčević's trichotomy that are not hereditarily separable. Together with these examples, Theorem P.29 determines which classes occur for convex response compacta, in the spirit of the Glasner–Megrelishvili hierarchy of tame systems [GM22].

Corollary P.31. *Every observable that is NIP$^{\mathrm{hs}}$ is NIP$^{\mathrm{met}}$.*

Proof. Let C be an observable on $H_1 \otimes H_2$ that is NIP$^{\mathrm{hs}}$. We show that the set $\overline{\mathcal{R}_C}$ of idealized responses satisfies the hypotheses of Theorem P.29. Since the slice map and evaluation at a probe are linear and $\mathfrak{D}(H_1)$ is convex, the identity $\mathfrak{r}_{t\rho+(1-t)\rho'} = t\,\mathfrak{r}_\rho + (1-t)\,\mathfrak{r}_{\rho'}$, for $\rho, \rho' \in \mathfrak{D}(H_1)$ and $0 \leq t \leq 1$, shows that the family $\mathcal{R}_C$ is convex. For fixed t, the map $(f, g) \mapsto tf + (1-t)g$ is continuous on $\mathbb{R}^{Y_C} \times \mathbb{R}^{Y_C}$ for the pointwise topology and sends $\mathcal{R}_C \times \mathcal{R}_C$ into $\mathcal{R}_C$; the map therefore sends $\overline{\mathcal{R}_C} \times \overline{\mathcal{R}_C}$, the closure of $\mathcal{R}_C \times \mathcal{R}_C$, into $\overline{\mathcal{R}_C}$, which is thus convex. By Lemma P.26 and Theorem 11.1(b), the set $\overline{\mathcal{R}_C}$ is a separable Rosenthal compactum, which is hereditarily separable because C is NIP$^{\mathrm{hs}}$. By Theorem P.29, the compactum $\overline{\mathcal{R}_C}$ is metrizable, that is, the observable C is NIP$^{\mathrm{met}}$. □

For observables, then, at most three of the four topological classes can occur, and in section P.2 we show that all three do. For families of continuous functions all four occur, as the minimal families show, and the class that observables miss is realized by D_3: the members of D_3, the characteristic functions of the sets $\{\zeta \in \{0,1\}^{\mathbb{N}} : \zeta \succeq s^\frown 0^\infty\}$ for s in the regular dyadic subtree Θ, are continuous on the Cantor set, since each of these sets is clopen, and their closure $\hat{S}_+(2^{\mathbb{N}})$ is hereditarily separable but not metrizable, so the family D_3 is $\mathrm{NIP}^{\mathrm{hs}}$ but not $\mathrm{NIP}^{\mathrm{met}}$.

This example comes from the lexicographic order of the Cantor set, which is isomorphic to the order that the real line induces on the middle-thirds Cantor set; the members of D_3 are realized types of this order, and they take only the values 0 and 1. Expectations bring in mixtures, and the midpoints of the split pairs turn the split interval into a space with an uncountable discrete subset; for the order of positions, the resulting space is the space of nondecreasing functions from $[0,1]$ to $[0,1]$ that take the value 0 at 0 and 1 at 1, a closed subspace of the Helly space of all nondecreasing functions from $[0,1]$ to $[0,1]$ (see section P.2). This is why, for averages over Alice's mixed preparations, hereditary separability forces metrizability (see Theorem P.29).

In the next theorem, increasing injections, equivalence of families and the minimal families D_5, D_6 and D_7 are those of section P.2.

Theorem P.32. *Let C be an observable on $H_1 \otimes H_2$ that is NIP for Alice's responses, and let $(\rho_n)_{n\in\mathbb{N}}$ be a trace-norm dense sequence in $\mathfrak{D}(H_1)$. Then exactly one of the following holds:*

(i) *the observable C is $\mathrm{NIP}^{\mathrm{met}}$;*

(ii) *the observable C is $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$, and there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that the responses $\{\mathfrak{r}_{\rho_{\iota(t)}} : t \in \{0,1\}^{<\mathbb{N}}\}$ are equivalent to D_6 or to D_7;*

(iii) *the observable C is not $\mathrm{NIP}^{\mathrm{fc}}$, and there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that the responses $\{\mathfrak{r}_{\rho_{\iota(t)}} : t \in \{0,1\}^{<\mathbb{N}}\}$ are equivalent to D_5.*

Proof. By Lemma P.22, the family $\{\mathfrak{r}_{\rho_n}\}$ is a pointwise bounded family of continuous functions on the compact metrizable space Y_C, and it is NIP because it is contained in $\mathcal{R}_C$. By Lemma P.26, the closure of this family is $\overline{\mathcal{R}_C}$. Hence Theorem P.20 applies, and case (ii) of that theorem is excluded by Corollary P.31. □

Physically, when the idealized responses of an NIP observable do not form a metrizable space, the reason is always visible in Alice's actual preparations: a family of them, arranged along the dyadic tree, has idealized limits that reproduce one of three canonical compacta, the closures of the minimal families D_5, D_6 and D_7. In the next section we locate the comparisons $[A < B]$ in this classification.

Three physical orders realize the three classes that occur for observables. Each of the three classes of Theorem P.32 contains a physically natural order, and the three orders are unstable and NIP. The comparison of photon numbers has a metrizable compactum of idealized responses and is thus $\mathrm{NIP}^{\mathrm{met}}$ (see Theorem P.40). The comparison of the positions of two particles, in a box or on a line, or of the quadratures of two modes of light (a quadrature is one of the two field amplitudes of a mode, which behave like a position and a momentum), has idealized responses that form a copy of the space of nondecreasing functions from $[0,1]$ to $[0,1]$ that take the value 0 at 0 and 1 at 1, a closed subspace of the Helly space; this space is first countable but not hereditarily separable, and the comparison is therefore $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$ (see Theorem P.40 and Proposition P.41). The causal order of a spacetime with one dimension of space, "Bob's detection lies in the causal future of Alice's event", is NIP

but not NIP$^{\mathrm{fc}}$: since a light cone can shrink to a point, the compactum of idealized responses of the causal order is not first countable (see Theorem P.42). The comparison of two phases, "Bob's angle is ahead of Alice's by less than half a turn", a cyclic relation rather than an order, falls in the same class, because on a circle the jumps at the two ends of a half-turn can cancel (see Theorem P.44).

Relations between classical quantities. All three orders relate a classical quantity measured by Alice to one measured by Bob, and their responses can be computed from classical probability distributions. Let (X_1, m_1) and (X_2, m_2) be σ-finite measure spaces such that the Hilbert spaces $H_1 = L^2(X_1, m_1)$ and $H_2 = L^2(X_2, m_2)$ are separable, fix an integer $d \geq 1$, let $\alpha\colon X_1 \to \mathbb{R}^d$ and $\beta\colon X_2 \to \mathbb{R}$ be measurable, and let $R \subseteq \mathbb{R}^d \times \mathbb{R}$ be a Borel set. The *relation observable* C_R is multiplication on $L^2(X_1 \times X_2, m_1 \otimes m_2) = H_1 \otimes H_2$ by the indicator of $\{(s,t) : (\alpha(s), \beta(t)) \in R\}$; the observable represents the event "Alice's value and Bob's value are related by R". In this section, the letters ρ and σ denote actual states of Alice's and of Bob's systems, that is, density operators on H_1 and on H_2. For a state ρ of Alice's system, the *distribution* of Alice's value in ρ is the Borel probability measure on $\mathbb{R}^d$ given by $\mu_\rho(U) = \operatorname{Tr}\big(\rho\, M_{\mathbf{1}_{\alpha^{-1}(U)}}\big)$ for Borel sets $U \subseteq \mathbb{R}^d$, where M_f denotes multiplication by f, and the distribution μ_σ of Bob's value in a state σ of Bob's system is defined in the same way. For $y \in \mathbb{R}$, let $R^y = \{x \in \mathbb{R}^d : (x, y) \in R\}$, and, for a state ρ of Alice's system, let $g_\rho\colon \mathbb{R} \to [0,1]$ be given by $g_\rho(y) = \mu_\rho(R^y)$, the probability that Alice's value is related by R to the value y; it is Borel measurable, by Tonelli's theorem. Recall from section P.2 that the effective observable $T_C(\rho)$ of an observable C is the observable on Bob's system obtained by fixing Alice's state ρ.

By the spectral theorem, every self-adjoint operator on a separable Hilbert space is unitarily equivalent to multiplication by a real measurable function on a space $L^2(X, m)$ of this kind. Thus, for self-adjoint operators A on H_1 and B on H_2, realized as multiplication by α and by β, the *comparison* $[A < B]$ is the relation observable C_R for $R = \{(x,y) : x < y\}$; it is the projection $(E_A \otimes E_B)(R)$, where E_A and E_B are the spectral measures, so the observable does not depend on the realization. When A and B are bounded, the operator $I \otimes B - A \otimes I$ is multiplication by $\beta(t) - \alpha(s)$, and $[A < B]$ is the spectral projection of this operator for the interval $(0, \infty)$, or equivalently the spectral projection of $A \otimes I - I \otimes B$ for the interval $(-\infty, 0)$.

Lemma P.33. *Let C_R be a relation observable, ρ a state of Alice's system and σ a state of Bob's system. Then the expectation value of C_R when the two systems are prepared independently in these states is*

$$\varphi_{C_R}(\rho, \sigma) = (\mu_\rho \otimes \mu_\sigma)(R) = \int g_\rho \, d\mu_\sigma,$$

and the effective observable $T_{C_R}(\rho)$ is multiplication by $g_\rho \circ \beta$.

Proof. For unit vectors $\psi \in H_1$ and $\chi \in H_2$, with distributions μ_ψ and μ_χ in the corresponding pure states, Tonelli's theorem and a change of variables give

$$\langle \psi \otimes \chi, C_R(\psi \otimes \chi)\rangle = \iint \mathbf{1}_R\big(\alpha(s), \beta(t)\big)\, |\psi(s)|^2 |\chi(t)|^2 \, dm_1(s)\, dm_2(t) = (\mu_\psi \otimes \mu_\chi)(R).$$

A density operator is a countable convex combination of pure states, $\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$, the series converging in trace norm; here j ranges over a countable set, the weights $\lambda_j \geq 0$ have sum 1, and $|\psi_j\rangle\langle\psi_j|$ is the orthogonal projection onto the line spanned by a unit vector $\psi_j \in H_1$. Accordingly, the distribution of Alice's value in ρ is $\mu_\rho = \sum_j \lambda_j \mu_{\psi_j}$. Both sides of the first identity are affine in each state and respect such countable combinations, so the first

identity holds for all states; the second is Tonelli's theorem. Finally, for every unit vector $\chi \in H_2$,

$$\langle \chi, T_{C_R}(\rho)\chi\rangle = \varphi_{C_R}(\rho, |\chi\rangle\langle\chi|) = \int g_\rho\big(\beta(t)\big)|\chi(t)|^2\, dm_2(t),$$

and a bounded operator on a complex Hilbert space is determined by these values. □

When the functions g_ρ are monotone, or differences of monotone functions, the responses cannot shatter large sets (see Proposition P.35 below). Recall from Remark P.25 that a finite set of responses is δ-shattered if every pattern on it is realized with reals $a < b$ such that $b - a \geq \delta$. For a real number $\delta > 0$, let $m_\delta = \lceil 9/\delta \rceil$, and let $N(\delta)$ be the largest positive integer n with $2^n \leq (2nm_\delta + 1)^{m_\delta}$; it exists because 2^n grows faster than any polynomial in n.

Lemma P.34. *Let $\delta > 0$ be a real number and n a positive integer, let $h_1, \dots, h_n \colon \mathbb{R} \to \mathbb{R}$ be functions of the form $h_i = u_i - v_i$, where $u_i, v_i \colon \mathbb{R} \to [0,1]$ are nondecreasing, and let $a < b$ be reals with $b - a \geq \delta$. Suppose that for every $S \subseteq \{1, \dots, n\}$ there is a Borel probability measure ν_S on $\mathbb{R}$ with $\int h_i\, d\nu_S \geq b$ for $i \in S$ and $\int h_i\, d\nu_S \leq a$ for $i \notin S$. Then $2^n \leq (2nm_\delta + 1)^{m_\delta}$, and hence $n \leq N(\delta)$.*

Proof. Write $m = m_\delta$. We first approximate a measure by m equally weighted points. Let ν be a Borel probability measure on $\mathbb{R}$, with distribution function $F(t) = \nu((-\infty, t])$. For $k = 1, \dots, m$, let $y_k = \min\{t : F(t) \geq (k - \frac{1}{2})/m\}$; the minimum exists because F is nondecreasing and right continuous, with limits 0 and 1 at $-\infty$ and $+\infty$, and $0 < (k - \frac{1}{2})/m < 1$. Let ν' be the average of the m point masses at the points $y_1, \dots, y_m$. Since $y_k \leq t$ exactly when $F(t) \geq (k - \frac{1}{2})/m$, we have $\nu'((-\infty, t]) = \lfloor mF(t) + \frac{1}{2} \rfloor / m$, whence $|\nu((-\infty, t]) - \nu'((-\infty, t])| \leq 1/m$ for every t; by taking limits from the left, we obtain the same bound for the half-lines $(-\infty, t)$. Hence $|\nu(J) - \nu'(J)| \leq 1/m$ for every set J of the form $\emptyset$, $\mathbb{R}$, $[t, \infty)$ or (t, ∞). For a nondecreasing $u \colon \mathbb{R} \to [0,1]$, every set $\{u > s\}$ has this form, and $\int u\, d\nu = \int_0^1 \nu(\{u > s\})\, ds$; therefore

$$\Big| \int u\, d\nu - \int u\, d\nu' \Big| \leq \frac{1}{m}. \tag{2}$$

Next we discretize the functions. Let $\tilde{u}_i = \lfloor mu_i \rfloor / m$ and $\tilde{v}_i = \lfloor mv_i \rfloor / m$. These functions are nondecreasing, take values in $\{0, \frac{1}{m}, \dots, 1\}$, and differ from u_i and v_i by at most $\frac{1}{m}$. Each of the $2n$ functions $\tilde{u}_i, \tilde{v}_i$ changes value at most m times along $\mathbb{R}$, so the map $Q(y) = (\tilde{u}_1(y), \dots, \tilde{u}_n(y), \tilde{v}_1(y), \dots, \tilde{v}_n(y))$ takes at most $2nm + 1$ values.

For each S, let $y_1^S, \dots, y_m^S$ be the points constructed from ν_S as above, and let $w(S) = (Q(y_1^S), \dots, Q(y_m^S))$. By (2), applied to u_i and to v_i, and by the bound on the discretization,

$$\Big| \int h_i\, d\nu_S - \frac{1}{m} \sum_{k=1}^{m} \big(\tilde{u}_i(y_k^S) - \tilde{v}_i(y_k^S)\big) \Big| \leq \frac{4}{m} \leq \frac{4\delta}{9},$$

and the average on the left is determined by $w(S)$. If $S \neq S'$, choose i in one of the two sets but not the other, say $i \in S \setminus S'$. Then $\int h_i\, d\nu_S - \int h_i\, d\nu_{S'} \geq \delta > \frac{8\delta}{9}$, so the two averages differ, and $w(S) \neq w(S')$. Thus the map $S \mapsto w(S)$ is injective, and $2^n \leq (2nm + 1)^m$. □

Proposition P.35. *Let C_R be a relation observable and $c \geq 1$ a constant such that, for every state ρ of Alice's system, $g_\rho = c\,(u_\rho - v_\rho)$ for two nondecreasing functions $u_\rho, v_\rho \colon \mathbb{R} \to [0,1]$. Then, for every $\delta > 0$, no set of more than $N(\delta/(2c))$ responses of Alice's states is δ-shattered by points of the probe space. Hence the observable C_R is uniformly NIP, with bounds that do not depend on the Hilbert spaces, and in particular NIP.*

PROOF. Suppose that the responses $\mathfrak{r}_{\rho_1}, \dots, \mathfrak{r}_{\rho_n}$ are δ-shattered with reals $a < b$, the pattern S being realized by a probe $\omega_{2,S}$ in the probe space Y_{C_R} of C_R. Let $a' = a + \delta/4$ and $b' = b - \delta/4$. The set of probes ω_2 with $\mathfrak{r}_{\rho_i}(\omega_2) > b'$ for $i \in S$ and $\mathfrak{r}_{\rho_i}(\omega_2) < a'$ for $i \notin S$ is open, by Lemma P.22(b), and contains $\omega_{2,S}$, so, by Lemma P.22(c), the set contains the record $\mathfrak{p}_{\sigma_S}$ of an actual state σ_S of Bob's system. By Lemma P.33, the value of $\mathfrak{r}_{\rho_i}$ there is $\int g_{\rho_i}\, d\mu_{\sigma_S}$, and Lemma P.34, applied with $h_i = g_{\rho_i}/c = u_{\rho_i} - v_{\rho_i}$, $\nu_S = \mu_{\sigma_S}$ and the reals $a'/c < b'/c$, yields $n \le N(\delta/(2c))$. An infinite independent sequence with reals $a < b$ would contain $(b-a)$-shattered sets of every finite size, so the observable C_R is NIP. □

Corollary P.36. *Let A and B be self-adjoint operators on H_1 and on H_2. Then, for every $\delta > 0$, no set of more than $N(\delta/2)$ responses of Alice's states for the comparison $[A < B]$ is δ-shattered by points of the probe space. In particular, the comparison $[A < B]$ is uniformly NIP, with bounds independent of the Hilbert spaces, and hence NIP.*

PROOF. Here $R^y = (-\infty, y)$, so the function $g_\rho(y) = \mu_\rho((-\infty, y))$ is nondecreasing, with values in $[0, 1]$. The hypothesis of Proposition P.35 therefore holds, with $c = 1$, $u_\rho = g_\rho$ and $v_\rho = 0$. □

Remark P.37 (Commuting observables and the general framework). The general results and the commuting examples play complementary roles. The definitions of section P.2 and the results of Theorem 11.1 and Theorem P.32 and Corollary P.31 apply to every bounded self-adjoint bipartite observable; for the swap test, for example, the probe algebra contains all compact operators. For every such observable the full compactum of idealized responses is convex; when the observable is NIP, hereditary separability therefore implies metrizability. The orders considered here have commutative probe algebras and realize all three resulting classes; the label-controlled observables of section P.2, in which Alice's label selects a test on Bob's system, also have commutative probe algebras. For these realizations, integration against classical probability distributions makes the responses explicit. For the relation observables that satisfy the monotonicity hypothesis of Proposition P.35, the proposition supplies bounds independent of the Hilbert spaces. In section P.2 we pass to parameterized magnetization families, including those of the Ising model in a transverse field, whose Hamiltonian has noncommuting terms; their classification follows from the same analysis of pointwise closures (see Theorem P.58).

Order observables. We now fix two model comparisons, one of discrete and one of continuous quantities, whose values lie in a compact set $\Xi \subseteq [0, 1]$ containing 0 and 1. The two cases are the following:

- *discrete*: the set Ξ is $\{1 - \frac{1}{l+1} : l \in \mathbb{N}\} \cup \{1\}$, the Hilbert spaces are $H_1 = H_2 = \ell^2(\mathbb{N})$, and the operator $A = B$ is the diagonal operator with eigenvalue $1 - \frac{1}{l+1}$ on the number ket $|l\rangle$; this operator is a bounded relabeling of the number operator, which counts photons and has eigenvalue l on $|l\rangle$, and the order of the eigenvalues is that of the photon numbers;
- *continuous*: the set Ξ is $[0, 1]$, the Hilbert spaces are $H_1 = H_2 = L^2[0, 1]$, and the operator $A = B$ is multiplication by x, the position of a particle confined to a box, measured from one end of the box in units of its length, so that it ranges over the interval $[0, 1]$.

In both cases, let $C = [A < B]$ be the spectral projection of $I \otimes B - A \otimes I$ for the interval $(0, \infty)$, which represents the event "Bob's value exceeds Alice's". For a state ρ, let μ_ρ be the distribution of A in ρ, a Borel probability measure on Ξ. In the discrete case every probability measure on $\Xi \setminus \{1\}$ occurs, and in the continuous case the measures that occur are exactly the absolutely continuous ones (for a probability density p, take the pure state given by the

unit vector $\psi = \sqrt{p}$). By Lemma P.33, the expectation value of C for Alice's state ρ and Bob's state σ is

$$\varphi_C(\rho,\sigma) = (\mu_\rho \otimes \mu_\sigma)\big(\{(x,y) : x < y\}\big) = \int g_\rho \, d\mu_\sigma, \qquad g_\rho(y) = \mu_\rho\big([0,y)\big).$$

The function $g_\rho\colon \Xi \to [0,1]$ is nondecreasing, with $g_\rho(0) = 0$ and $g_\rho(1) = 1$ in both cases, since μ_ρ assigns no mass to the point 1; in the continuous case the map is also continuous. Accordingly, the effective observable is $T_C(\rho) = g_\rho(B)$. By Corollary P.36, the observable C is uniformly NIP, and in particular NIP. It is unstable, by the interleaving argument of Proposition 2.2, which uses only the order of the values. In the discrete case, let Alice prepare the number states $\rho_k = |2k\rangle\langle 2k|$ and Bob the number states $\sigma_l = |2l+1\rangle\langle 2l+1|$; in the continuous case, let ρ_k and σ_l be states of Alice's and of Bob's particles supported in small disjoint intervals around positions a_k and b_l, respectively, for $k, l \geq 1$, interleaved as $a_1 < b_1 < a_2 < b_2 < \cdots$. In both cases we obtain $\varphi_C(\rho_k, \sigma_l) = 1$ for $k \leq l$ and 0 for $k > l$; the iterated limit taken first in k is 0, whereas the one taken first in l is 1.

The probe space and the idealized responses. For both comparisons of the previous subsection, the discrete and the continuous one, and for every state ρ of Alice's system, the function g_ρ, whose value at a point $y \in \Xi$ is the probability that Alice's value lies below y, is nondecreasing, and the pointwise limits of these functions describe all idealized responses (see Proposition P.39 below). Recall that the probe algebra $\mathfrak{A}_C$ is the unital C*-algebra generated by the effective observables, and that the probe space Y_C is the state space. We write $C(\Xi)$ for the algebra of continuous complex functions on Ξ, with the supremum norm, and M_Ξ for the set of nondecreasing functions $g\colon \Xi \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$, with the topology of pointwise convergence; the set M_Ξ is a closed subset of $[0,1]^\Xi$, hence compact. Although M_Ξ contains discontinuous functions, integration against a fixed probability measure is continuous on it, by the following lemma; on the compact metrizable space Ξ, Radon and Borel probability measures coincide, so the lemma applies to every point of the probe space described in Proposition P.39.

Lemma P.38. *Let ν be a Borel probability measure on Ξ, and let (g_i) be a net in M_Ξ, indexed by a directed set, that converges pointwise to a function $g \in M_\Xi$. Then $\int g_i \, d\nu \to \int g \, d\nu$.*

Proof. Let $\varepsilon > 0$ be given. Choose a positive integer q with $1/q < \varepsilon$. For each $j = 1, \ldots, q-1$, take the least point $x \in \Xi$ with $\nu(\Xi \cap [0,x]) \geq j/q$, and adjoin 0 and 1. List the distinct points obtained as $0 = p_0 < p_1 < \cdots < p_m = 1$. Each open cell $c_k = \Xi \cap (p_{k-1}, p_k)$ has measure at most $1/q < \varepsilon$: a cell of greater measure would contain a point at which one of the intervening levels j/q is first reached, contrary to the choice of consecutive partition points. Let $\eta_i = \max_k |g_i(p_k) - g(p_k)|$, which tends to 0. On c_k both g_i and g take values between their values at p_{k-1} and at p_k, because they are nondecreasing. Hence

$$\Big| \int_{c_k} g_i \, d\nu - \int_{c_k} g \, d\nu \Big| \leq \nu(c_k)\big(g(p_k) - g(p_{k-1}) + 2\eta_i\big),$$

and summing over k, together with the contribution of the points p_k, shows that $|\int g_i \, d\nu - \int g \, d\nu| \leq \varepsilon + 3\eta_i$, since the differences $g(p_k) - g(p_{k-1})$ add up to at most 1. □

Proposition P.39. *For each of the two comparisons $C = [A < B]$, the discrete and the continuous one, the probe algebra $\mathfrak{A}_C$ is $\{f(B) : f \in C(\Xi)\}$, so the state space Y_C is identified with the space $P(\Xi)$ of Radon probability measures on Ξ, with the weak* topology, by $\nu \mapsto [f(B) \mapsto \int_\Xi f \, d\nu]$. Under this identification, the record $\mathfrak{p}_\sigma$ of an actual state σ of Bob's system corresponds to μ_σ: these are exactly the measures with $\mu_\sigma(\{1\}) = 0$ in the discrete case and*

the absolutely continuous measures in the continuous case. The map $\Phi\colon \mathfrak{r} \mapsto \big(\mathfrak{r}(\delta_y)\big)_{y\in\Xi}$, *where* δ_y *is the point mass at* y, *is a homeomorphism from the compactum* $\overline{\mathcal{R}_C}$ *of idealized responses onto* M_Ξ, *and every idealized response is given by* $\mathfrak{r}(\nu) = \int \Phi(\mathfrak{r})\, d\nu$.

Proof. Each function g_ρ is continuous on Ξ: in the continuous case this follows from absolute continuity of the distribution, and in the discrete case continuity at the only accumulation point follows from $g_\rho(1 - 1/(l+1)) \to 1 = g_\rho(1)$ as $l \to \infty$. The functions g_ρ for states ρ whose distribution has full support in Ξ separate the points of Ξ, and the map $f \mapsto f(B)$ is an isometric $*$-isomorphism from $C(\Xi)$ onto the image, because the spectrum of B is Ξ. By the Stone–Weierstrass theorem, the unital C*-algebra generated by the operators $g_\rho(B)$ is $\{f(B) : f \in C(\Xi)\}$, and the slice map takes values in it. By the Riesz representation theorem, the state space of this algebra is $P(\Xi)$, and the response of ρ is $\mathfrak{r}_\rho(\nu) = \int g_\rho\, d\nu$.

The map Φ is continuous, because the coordinates of Φ are evaluations. Let $\mathfrak{r} \in \overline{\mathcal{R}_C}$ be the pointwise limit of a net $(\mathfrak{r}_{\rho_i})$. Then the functions $g_{\rho_i} = \Phi(\mathfrak{r}_{\rho_i})$ converge pointwise to $g = \Phi(\mathfrak{r}) \in M_\Xi$, and, by Lemma P.38, $\mathfrak{r}(\nu) = \lim_i \int g_{\rho_i}\, d\nu = \int g\, d\nu$. Thus the response $\mathfrak{r}$ is determined by $\Phi(\mathfrak{r})$, the map Φ is injective, and, as a continuous injection from a compact space, the map Φ is a homeomorphism onto the image. The image is compact and contains the functions g_ρ.

It remains to be shown that the functions g_ρ are dense in M_Ξ. Let $g \in M_\Xi$ and a finite set $F \subseteq \Xi$ containing 0 and 1 be given. In the continuous case, the piecewise linear interpolation $\tilde{g}$ of g on F is continuous and nondecreasing, with values 0 at 0 and 1 at 1; hence $\tilde{g} = g_\rho$ for a state ρ whose distribution has $\tilde{g}$ as distribution function, which is absolutely continuous, and $\tilde{g}$ agrees with g on F. In the discrete case, write $F = \{\xi_0 < \xi_1 < \cdots < \xi_m\}$, with $\xi_0 = 0$ and $\xi_m = 1$. The probability measure with mass $g(\xi_{j+1}) - g(\xi_j)$ at ξ_j for $0 \le j < m$ is the distribution of a state ρ with finitely many photons, and $g_\rho(\xi_j) = g(\xi_j) - g(\xi_0) = g(\xi_j)$ for every j, so the function g_ρ agrees with g on F. Every neighborhood of g therefore contains some g_ρ. □

Location in the classification. We can now place both comparisons in the classification of Theorem P.32. Recall from section P.2 that D_6 and D_7 are the two minimal families of the heptachotomy whose closures are first countable but not hereditarily separable.

Theorem P.40. *Let* $C = [A < B]$ *be the comparison of Alice's and Bob's photon numbers, in the discrete case, or of the positions of their particles in a box, in the continuous case.*

(a) *In the discrete case, the observable* C *is unstable and* NIP$^{\mathrm{met}}$.

(b) *In the continuous case, the observable* C *is unstable and* NIP$^{\mathrm{fc}}$, *but not* NIP$^{\mathrm{hs}}$. *Hence, for every trace-norm dense sequence* $(\rho_n)_{n\in\mathbb{N}}$ *of states of Alice's particle, there is an increasing injection* $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ *such that the responses* $\{\mathfrak{r}_{\rho_{\iota(t)}} : t \in \{0,1\}^{<\mathbb{N}}\}$ *are equivalent to* D_6 *or to* D_7.

Proof. (a) By Proposition P.39, the compactum of idealized responses is homeomorphic to a subspace of $[0,1]^\Xi$ with Ξ countable, which is metrizable.

(b) By Proposition P.39, the compactum of idealized responses is homeomorphic to $M_{[0,1]}$, a copy of the closed subspace of the Helly space described at the start of this section. It is first countable. Given $g \in M_{[0,1]}$, let Z_g be the countable set consisting of the rational points, the jump points of g and the endpoints; we claim that the sets $\{h : |h(p) - g(p)| < 1/n \text{ for } p \in F\}$, with $F \subseteq Z_g$ finite and $n \ge 1$, form a neighborhood base at g. It suffices to show that each set $\{h : |h(y) - g(y)| < \varepsilon\}$, with $y \in [0,1]$ and $\varepsilon > 0$, contains one of them. If $y \in Z_g$, this is clear. If $y \notin Z_g$, then g is continuous at y, and there are rationals $p < y < q$ with

$g(q)-g(p)<\varepsilon/2$. Every $h\in M_{[0,1]}$ with $|h-g|<\varepsilon/4$ at p and q satisfies $h(p)\le h(y)\le h(q)$, because h is nondecreasing, and hence $|h(y)-g(y)|<\varepsilon$.

The compactum is not hereditarily separable. For $0<a<1$, let h_a be 0 on $[0,a)$, $\frac12$ at a and 1 on $(a,1]$. The set $\{h\in M_{[0,1]}:\frac14<h(a)<\frac34\}$ is a neighborhood of h_a that contains no other h_b, so the set $\{h_a:0<a<1\}$ is an uncountable discrete subspace.

The observable is therefore $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$, and the last statement is Theorem P.32(ii). □

The idealized responses of the continuous comparison have a direct physical reading, which is also a probabilistic one. They are described by all the functions g in $M_{[0,1]}$, where $g(y)$ is the limiting probability that Alice's value lies below Bob's fixed value y. Such a function need not be the distribution function of a single countably additive position measure: the value of the response at a jump also records how the particle approaches the jump point, from the left, from the right, or both. The function h_a describes a half-and-half mixture of a particle concentrating at a from the left and one concentrating at a from the right, and Bob's idealized probe concentrated at a reads the value $\frac12$. In the discrete case no such splitting occurs at an eigenvalue, because each eigenvalue is isolated; the additional limiting behavior is escape to infinite photon number, the point 1 of Ξ, which may be mixed with an ordinary distribution on the eigenvalues.

Unconfined quantities. Confinement to a box is not what produces this subspace of the Helly space: a comparison depends only on the order of the two outcomes, so the comparison is unchanged when both quantities are replaced by the same strictly increasing function of themselves, and a suitable function turns unconfined quantities, such as the positions of two particles on a line or the quadratures of two modes of light, into positions in a box (see Proposition P.41 below).

Proposition P.41. *Let A be a self-adjoint operator on the Hilbert space H_1 of Alice's system and B one on the Hilbert space H_2 of Bob's, each unitarily equivalent to multiplication by x on $L^2(\mathbb{R})$. Then the comparison $[A<B]$ is unstable, and the idealized responses of the comparison form a copy of the space of nondecreasing functions $[0,1]\to[0,1]$ taking the value 0 at 0 and 1 at 1, a closed subspace of the Helly space. Hence the comparison is $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$.*

Proof. We may assume that A and B are multiplication by x on $L^2(\mathbb{R})$. For physical quantities, x denotes the numerical value in the fixed common unit used for the comparison. Let $\theta\colon\mathbb{R}\to(0,1)$ be the increasing diffeomorphism $\theta(x)=\frac12(1+\tanh x)$. The operator U given by $(U\psi)(u)=\psi(\theta^{-1}(u))\,|(\theta^{-1})'(u)|^{1/2}$ is a unitary from $L^2(\mathbb{R})$ onto $L^2[0,1]$ that carries multiplication by $\theta(x)$ to multiplication by u. Since θ is strictly increasing, $x<y$ exactly when $\theta(x)<\theta(y)$; hence the spectral projection $[A<B]$ equals $[\theta(A)<\theta(B)]$, and $U\otimes U$ carries this projection to the comparison of the continuous case. The unitary $U\otimes U$ preserves expectation values and carries actual and idealized states to actual and idealized states; the unitary therefore induces a homeomorphism between the two compacta of idealized responses. The statement now follows from Proposition P.39 and Theorem P.40(b). □

For the comparisons covered by Proposition P.41, confinement does not affect the compactum of idealized responses. For quantities whose expectation values matter, such as energies, the rescaling by θ is not available, because the rescaling is not affine.

The causal order is not $\mathrm{NIP}^{\mathrm{fc}}$. The compactum of idealized responses of a comparison is first countable because the members of the compactum are monotone, and a monotone

function has at most countably many jumps, each with two sides (see the proof of Theorem P.40(b)). A relation between Bob's value and two values of Alice's, one on each side, removes this restriction (see Theorem P.42 below), and the natural example is the causal order of spacetime. We work in a $(1+1)$-dimensional spacetime, with one dimension of time and one of space, in units in which the speed of light is 1. An event (t, ξ), at time t and position ξ, can influence a detection by Bob at the point y at time 0 exactly when $|y - \xi| \le -t$, that is, when y lies between the points $a = \xi + t$ and $b = \xi - t$ at which the two light rays through the event cross Bob's time slice. We describe Alice's event by these null coordinates and confine everything to the causal diamond over Bob's interval $[0, 1]$, the set of events whose two light rays cross Bob's slice inside the interval: Alice's Hilbert space is $H_1 = L^2([0,1]^2)$, with variables (a, b), Bob's is $H_2 = L^2[0, 1]$, with variable y, and the *causal-order observable* C is multiplication by the indicator of

$$R = \{(a, b, y) : a \le y \le b\}.$$

Events with $a > b$ lie in the future of Bob's slice and influence none of his detections. Recall from section P.2 that D_5 is the minimal family of the heptachotomy whose closure, $\hat{A}(2^{\mathbb{N}})$, is not first countable.

For a state ρ of Alice's system, let μ_ρ be the distribution of (a, b) in ρ, an absolutely continuous probability measure on $[0,1]^2$, and similarly for Bob. The observable C is the relation observable C_R for this relation, so, by Lemma P.33, the expectation value of the observable in states ρ and σ of the two systems is $\varphi_C(\rho, \sigma) = (\mu_\rho \otimes \mu_\sigma)(R) = \int g_\rho \, d\mu_\sigma$, where $g_\rho(y) = \mu_\rho(\{(a, b) : a \le y \le b\})$ is the probability that Alice's event can influence a detection at y. The function g_ρ is continuous: the sections of R at y and at y' differ by a set contained in two strips, each of area at most $|y - y'|$, and the μ_ρ-measure of such strips tends to 0 with $|y - y'|$, because μ_ρ is absolutely continuous. The effective observable $T_C(\rho)$ is multiplication by g_ρ.

Theorem P.42. *The causal-order observable is NIP but not NIP*$^{\mathrm{fc}}$*. Hence, for every trace-norm dense sequence $(\rho_n)_{n\in\mathbb{N}}$ of states of Alice's system, there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that the responses $\{\mathfrak{r}_{\rho_{\iota(t)}} : t \in \{0,1\}^{<\mathbb{N}}\}$ are equivalent to D_5.*

Proof. For every state ρ, the event $\{a \le y \le b\}$ is the event $\{a \le y\}$ with the event $\{a \le y,\ b < y\}$ removed, so

$$g_\rho(y) = \mu_\rho\big(\{(a, b) : a \le y\}\big) - \mu_\rho\big(\{(a, b) : a \le y,\ b < y\}\big),$$

a difference of two nondecreasing functions with values in $[0, 1]$. By Proposition P.35, whose constant is here 1, the observable C is uniformly NIP, and in particular NIP. By Lemma P.26 the compactum of idealized responses is separable, so the level of the observable is defined.

The probe algebra consists of operators M_f of multiplication by continuous functions f; the probe algebra is therefore a C*-subalgebra of $\{M_f : f \in C[0, 1]\}$, which is isometrically isomorphic to $C[0, 1]$, because $\|M_f\| = \sup |f|$. Every state of $\mathfrak{A}_C$ extends to a state of this larger algebra (see [**KR83**, Theorem 4.3.13]), and, by the Riesz representation theorem, every state ω_2 of $\mathfrak{A}_C$ is thus represented by a probability measure ν on $[0, 1]$, in the sense that $\omega_2(M_f) = \int f \, d\nu$ for $M_f \in \mathfrak{A}_C$; in particular, $\mathfrak{r}_\rho(\omega_2) = \int g_\rho \, d\nu$. For each point $c \in (0, 1)$, the evaluation $\omega_{2,c}(M_f) = f(c)$ is a state of $\mathfrak{A}_C$, Bob's probe localized at c.

Let ρ_0 be a state whose distribution is uniform on $[\frac{1}{2}, 1] \times [0, \frac{1}{2}]$, an event after Bob's slice. Then $g_{\rho_0} = 0$, and the zero function is a response. For $c \in (0, 1)$ and large n, let $\rho_{c,n}$ be a state whose distribution is uniform on $[c - \frac{2}{n}, c - \frac{1}{n}] \times [c + \frac{1}{n}, c + \frac{2}{n}]$, an event at position near c shortly before Bob's slice. Then $g_{\rho_{c,n}}(c) = 1$, and $g_{\rho_{c,n}}(y) = 0$ when $|y - c| > \frac{2}{n}$. Hence the

function $g_{\rho_{c,n}}$ converges pointwise to the indicator of $\{c\}$, and, by dominated convergence, $\mathfrak{r}_{\rho_{c,n}}(\omega_2) \to \nu(\{c\})$ for every probe ω_2 and every measure ν representing it. The limit $\mathfrak{r}_c$ is an idealized response, with $\mathfrak{r}_c(\omega_{2,c}) = 1$.

Suppose that the zero function had a countable base of neighborhoods (U_k) in the compactum of idealized responses. Each U_k contains a set $\{\mathfrak{r} : |\mathfrak{r}(\omega_2)| < \varepsilon_k \text{ for } \omega_2 \in F_k\}$ with F_k finite. Choose a representing measure for each probe in $F = \bigcup_k F_k$; together these measures have countably many atoms, so some $c \in (0,1)$ is an atom of none of them. Then $\mathfrak{r}_c(\omega_2) = 0$ for $\omega_2 \in F$, so the response $\mathfrak{r}_c$ lies in every U_k. Since the space is Hausdorff, the intersection of a neighborhood base at a point is the singleton of the point; hence $\mathfrak{r}_c = 0$, contrary to $\mathfrak{r}_c(\omega_{2,c}) = 1$. The compactum is therefore not first countable, and the last statement is Theorem P.32(iii). □

Physically, the proof says the following. For $c \in (0,1)$, the idealized response $\mathfrak{r}_c$ is the limit of responses of preparations whose event coordinates concentrate near $(a,b) = (c,c)$ from $a < c < b$. The corresponding light-cone intervals shrink to c, so only probes whose measures have an atom at c detect the limiting response. Every actual probe has an absolutely continuous position distribution and therefore gives $\mathfrak{r}_c$ the value 0; the idealized probe $\omega_{2,c}$ instead gives the value 1. There is one such idealized response for each point c in this interval, and no countable family of probes distinguishes all of them from the zero response, as the neighborhood-base argument shows. The relation R is the intersection of two comparisons, $a \le y$ and $y \le b$, one for each light ray. A single comparison has monotone idealized responses; the two rays together allow idealized responses supported on a single point. Here Todorčević's trichotomy is used in full: the theorem turns the failure of first countability into a dyadic tree of actual preparations of Alice's event coordinates whose responses are equivalent to the minimal family D_5. The proof suggests a tree of this kind, namely preparations localized near events whose light cones meet Bob's slice in the intervals of a Cantor construction, each event later than its predecessor, although we do not claim that the trichotomy produces this one.

The comparison of phases is not $\mathrm{NIP}^{\mathrm{fc}}$. Orders on a line give first countable compacta of idealized responses (see Theorem P.40 and Proposition P.41); on a circle, first countability can fail (see Theorem P.44 below). Identify the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with $[0,1)$, so that angles are measured in turns and a half-turn is $\frac{1}{2}$. Alice and Bob each hold a particle on a ring, a planar rotor, with Hilbert spaces $H_1 = H_2 = L^2(\mathbb{T})$, the angle acting by multiplication. The *phase comparison* is the relation observable $C = C_R$ for

$$R = \big\{(x,y) \in [0,1)^2 : \ y - x \in (0, \tfrac{1}{2}) \bmod 1\big\},$$

the event "Bob's angle is ahead of Alice's by less than half a turn". For a state ρ of Alice's rotor, the distribution μ_ρ of her angle is an absolutely continuous probability measure on $\mathbb{T}$, and every such measure occurs. By Lemma P.33, the phase comparison has, in states ρ and σ of the two rotors, the expectation value $\varphi_C(\rho,\sigma) = \int g_\rho \, d\mu_\sigma$, where $g_\rho(y) = \mu_\rho\big((y - \frac{1}{2}, y)\big)$ is the probability that Alice's angle lies in the open half-circle ending at y; the function g_ρ is continuous on $\mathbb{T}$, because μ_ρ has no atoms.

Lemma P.43. *For every probability measure μ on $\mathbb{T} = [0,1)$, the function $g(y) = \mu((y - \frac{1}{2}, y))$ on $[0,1)$ is $2(u - v)$ for two nondecreasing functions $u, v \colon [0,1) \to [0,1]$.*

Proof. Let $G(y) = \mu([0,y))$. For $y \in [\frac{1}{2}, 1)$ the half-circle is the interval $(y - \frac{1}{2}, y)$, and $g(y) = G(y) - \mu([0, y - \frac{1}{2}])$; for $y \in [0, \frac{1}{2})$ the set is $(y + \frac{1}{2}, 1) \cup [0, y)$, and $g(y) = G(y) + 1 - \mu([0, y + \frac{1}{2}])$. Let $V(y) = \mu([0, y + \frac{1}{2}]) - 1$ for $0 \le y < \frac{1}{2}$ and $V(y) = \mu([0, y - \frac{1}{2}])$

for $\frac{1}{2} \le y < 1$. Then $g = G - V$. The functions G and V are nondecreasing, with values in $[0,1]$ and $[-1,1]$, respectively; indeed, the function V is nondecreasing on each half, and at $\frac{1}{2}$ the function V passes from $\lim_{y\uparrow 1/2} V(y) = 0$ to $V(\frac{1}{2}) = \mu(\{0\}) \ge 0$. Let $u = \frac{1}{2}(G+1)$ and $v = \frac{1}{2}(V+1)$; then $g = 2(u-v)$. □

Theorem P.44. *For every $\delta > 0$, no set of more than $N(\delta/4)$ responses of Alice's states for the phase comparison is δ-shattered by points of the probe space, so the phase comparison is uniformly NIP; it is not NIP$^{\mathrm{fc}}$. Hence, for every trace-norm dense sequence $(\rho_n)_{n\in\mathbb{N}}$ of states of Alice's rotor, there is an increasing injection $\iota\colon \{0,1\}^{<\mathbb{N}} \to \mathbb{N}$ such that the responses $\{\mathfrak{r}_{\rho_{\iota(t)}} : t \in \{0,1\}^{<\mathbb{N}}\}$ are equivalent to D_5.*

Proof. By Lemma P.43, the function g_ρ agrees on $[0,1)$ with $2(u-v)$ for two nondecreasing functions $u, v\colon [0,1) \to [0,1]$, which we extend to nondecreasing functions on $\mathbb{R}$, constant on $(-\infty, 0)$ and on $[1,\infty)$. The function g_ρ vanishes off $[0,1)$, where the extended difference $2(u-v)$ need not vanish, so the hypothesis of Proposition P.35 holds only on $[0,1)$. Bob's distributions are carried by $[0,1)$, however, and only these values of g_ρ enter the proof of the proposition; the proof therefore applies verbatim with the constant 2, which gives the bound $N(\delta/4)$.

As in the proof of Theorem P.42, every probe ω_2 is represented by a probability measure ν on $\mathbb{T}$, with $\mathfrak{r}_\rho(\omega_2) = \int g_\rho \, d\nu$, and for $z \in \mathbb{T}$ the evaluation $\omega_{2,z}(M_f) = f(z)$ is an idealized probe localized at z. An actual state σ instead induces the probe represented by the absolutely continuous angle distribution μ_σ. Let ρ_0 be a state in which Alice's angle is uniformly distributed. Every open half-circle has measure $\frac{1}{2}$, so $\mathfrak{r}_{\rho_0}(\omega_2) = \frac{1}{2}$ for every probe ω_2.

For $a \in \mathbb{T}$ and large n, let $\rho_{a,n}$ be a state whose distribution is uniform, with weight $\frac{1}{2}$ each, on the arcs $\Delta_1 = (a - \frac{1}{n}, a)$ and $\Delta_2 = (a + \frac{1}{2}, a + \frac{1}{2} + \frac{1}{n})$: half of the distribution lies just behind a, half just ahead of the antipode. Write $A_y = (y - \frac{1}{2}, y)$. If $y = a$, then the half-circle A_y contains Δ_1 and Δ_2, so $g_{\rho_{a,n}}(a) = 1$. If $y = a + \frac{1}{2}$, then the half-circle $A_y = (a, a + \frac{1}{2})$ meets neither, and $g_{\rho_{a,n}}(a + \frac{1}{2}) = 0$. Otherwise the endpoints of A_y are at some positive distance d from a and from $a + \frac{1}{2}$; an open half-circle whose endpoints differ from these two antipodal points contains exactly one of them. Once $\frac{1}{n} < d$, the half-circle A_y therefore contains the arc Δ_1 just behind a if A_y contains a, and the arc Δ_2 just ahead of $a + \frac{1}{2}$ if it contains $a + \frac{1}{2}$, but not both. Thus $g_{\rho_{a,n}}(y) = \frac{1}{2}$. Hence the function $g_{\rho_{a,n}}$ converges pointwise to $g_a = \frac{1}{2} + \frac{1}{2}\mathbf{1}_{\{a\}} - \frac{1}{2}\mathbf{1}_{\{a+1/2\}}$, and, by dominated convergence, $\mathfrak{r}_{\rho_{a,n}}(\omega_2) \to \frac{1}{2} + \frac{1}{2}\nu(\{a\}) - \frac{1}{2}\nu(\{a + \frac{1}{2}\})$ for every probe ω_2 and every measure ν representing it. The limit $\mathfrak{r}_a$ is an idealized response, with $\mathfrak{r}_a(\omega_{2,a}) = 1$.

Suppose that $\mathfrak{r}_{\rho_0}$ had a countable base of neighborhoods (U_k) in the compactum of idealized responses. Each U_k contains a set $\{\mathfrak{r} : |\mathfrak{r}(\omega_2) - \frac{1}{2}| < \varepsilon_k \text{ for } \omega_2 \in F_k\}$ with F_k finite. Choose a representing measure for each probe in $F = \bigcup_k F_k$; together these measures have countably many atoms, so there is $a \in \mathbb{T}$ such that neither a nor $a + \frac{1}{2}$ is one of them. Then $\mathfrak{r}_a(\omega_2) = \frac{1}{2}$ for every $\omega_2 \in F$, so $\mathfrak{r}_a$ lies in every U_k, and, the space being Hausdorff, the response $\mathfrak{r}_a$ equals $\mathfrak{r}_{\rho_0}$, contrary to $\mathfrak{r}_a(\omega_{2,a}) = 1$. The compactum is therefore not first countable, and the last statement is Theorem P.32(iii). □

Here, too, the proof has a physical reading. A state in which the angle is uniformly distributed responds $\frac{1}{2}$ to every probe. For an antipodal pair of points a and $a + \frac{1}{2}$, the idealized limit of preparations that place half of the angle's distribution just behind a and half just ahead of $a + \frac{1}{2}$ also responds $\frac{1}{2}$, except that a probe at a reads 1 and a probe at the antipode reads 0: the jumps at the two ends of the half-turn point in opposite directions

and cancel everywhere else. There is one such pair for each point of the circle, and countably many probes cannot distinguish all the resulting responses from that of the uniform state. On a line this cannot happen, because a monotone response cannot hide a jump (see the proof of Theorem P.40(b)); on a circle there is no first or last point. The same contrast appears in topological dynamics. The enveloping semigroup of a dynamical system is the pointwise closure of the acting maps of the system. Glasner and Megrelishvili showed that compact metrizable linearly ordered spaces, with their order topologies and order-preserving group actions, have first countable enveloping semigroups, while the circular analogue of the Helly space is not first countable [**GM22**, Theorem 8.8, Proposition 8.10 and Remark 8.11]. The circular space in their comparison consists of circular-order-preserving maps of the circle, whereas the responses here are real-valued; neither result implies the other.

Interference fringes are not NIP.. The relations studied so far in this section are all NIP. The fixed-contrast interference patterns in Proposition P.45 below show how oscillation at ever finer scales can produce the independence property. Positions on Bob's screen are measured in a fixed unit of length, in which the screen is the interval $[0, 2\pi]$, and wavenumbers are measured in the inverse unit. We model the experiment by a register on Alice's side with orthonormal basis $(e_n)_{n\in\mathbb{N}}$, whose label n selects a fringe pattern with wavenumber $k_n = 8^n$. At position $x \in [0, 2\pi]$, Bob's detector accepts with probability $f_n(x) = \frac{1}{2}(1 + \cos k_n x)$, a model of the local detector response that is largest on a bright fringe. The *fringe observable* is $C = \sum_n |e_n\rangle\langle e_n| \otimes M_{f_n}$ on $\ell^2(\mathbb{N}) \otimes L^2([0, 2\pi], \lambda)$, where λ is Lebesgue measure divided by 2π: the label n of Alice's state selects the test f_n applied to Bob's detector, and C is thus a label-controlled observable in the sense of Definition P.46 below.

Proposition P.45. *The fringe observable is not NIP: Alice's basis states e_n form an independent sequence, and the closure of their responses contains a copy of $\beta\mathbb{N}$.*

Proof. Let $E, F \subseteq \mathbb{N}$ be finite and disjoint. Choose $y \in [0, 1)$ whose base-8 digit in place $n + 1$ is 0 for $n \in F$ and 4 for $n \in E$, and let $x = 2\pi y$. Then $\cos k_n x = \cos(2\pi\{8^n y\})$, where $\{8^n y\}$, the fractional part of $8^n y$, lies in $[0, \frac{1}{8})$ for $n \in F$ and in $[\frac{1}{2}, \frac{5}{8})$ for $n \in E$; hence $f_n(x) \geq \frac{3}{4}$ for $n \in F$ and $f_n(x) \leq \frac{1}{4}$ for $n \in E$. By continuity of the finitely many functions f_n involved, a state σ of Bob's detector concentrated near x has expectation values at least $\frac{5}{8}$ for $n \in F$ and at most $\frac{3}{8}$ for $n \in E$. By Lemma P.22(c), the restrictions of such states are probes; the sequence (e_n) is therefore independent. As in the proof of Theorem 11.1(c), no subsequence of the responses converges pointwise, and Fact P.9 yields a subsequence whose closure is homeomorphic to $\beta\mathbb{N}$. □

The contrast between these fringes and the orders of this section displays the finite-pattern obstruction measured by NIP. A comparison asks on which side of a threshold a value lies; the pattern of wavenumber 8^n has 8^n periods on Bob's screen, each with a bright and a dark half. The base-8 construction lets one prescribe bright or dark responses independently for any finite collection of labels. This is independence of finite patterns, rather than statistical independence of the fringe intensities under a distribution of screen positions.

The seven minimal families in physics. The classification of a full response compactum and the classification of a selected tree of preparations concern different compact sets. For an NIP observable, the full compactum is convex, and hereditary separability implies metrizability by Corollary P.31; the closure of a selected tree need not be convex. Each of the seven minimal families is nevertheless realized by such a tree of actual preparations (see

Theorem P.50 below). Photon-number comparisons, position comparisons and spin registers provide the realizations, while the passage from functions of position to responses on probability measures preserves the indexed families and their pointwise closures.

Label-controlled observables. Several of our observables are of the following kind. Alice's system is a register that holds a label, and the label selects the test applied to Bob's system: a detector that accepts Bob's system at each point with a probability that depends on the label and varies continuously with the point.

Definition P.46. Let L be a compact metrizable space, λ a Borel probability measure on L with full support, I a countable set and $(f_i)_{i\in I}$ a family of continuous functions $L \to [0,1]$. The *label-controlled observable* is

$$C = \sum_{i\in I} |e_i\rangle\langle e_i| \otimes M_{f_i} \qquad \text{on } \ \ell^2(I) \otimes L^2(L,\lambda),$$

where (e_i) is the standard basis of $\ell^2(I)$ and M_f is multiplication by f.

The sum converges strongly, that is, on each vector, because the terms act on orthogonal subspaces, and the operator C lies between 0 and the identity operator. When L contains a countable dense set Z of isolated points, we take $\lambda = \sum_{x\in Z} w_x \delta_x$, a combination of the point masses δ_x with positive weights w_x of sum 1; then the space $L^2(L,\lambda)$ is a weighted copy of $\ell^2(Z)$, and the position of Bob's system takes values in Z. As in section P.2, we write $C(L)$ for the algebra of continuous complex functions on L, with the supremum norm, and $P(L)$ for the space of Borel probability measures on L, with the weak* topology. Recall that the probe algebra $\mathfrak{A}_C$ of an observable C is the unital C*-algebra generated by the effective observables of C, and that the probe space Y_C is the state space of $\mathfrak{A}_C$.

Lemma P.47. *Let C be as in Definition P.46, and suppose that the functions f_i separate the points of L. Then the probe algebra $\mathfrak{A}_C$ is $\{M_f : f \in C(L)\}$, the probe space Y_C is identified with $P(L)$, and the response of the label state e_i is the function $\mathfrak{r}_{e_i} \colon P(L) \to [0,1]$ given by $\mathfrak{r}_{e_i}(\nu) = \int f_i \, d\nu$. The record $\mathfrak{p}_\sigma$ of an actual state σ of Bob's system corresponds to the measure μ_σ characterized by $\operatorname{Tr}(\sigma M_f) = \int f \, d\mu_\sigma$ for $f \in C(L)$. The measures obtained in this way are exactly those absolutely continuous with respect to λ.*

Proof. For a density operator ρ, tracing out Alice's system yields the norm-convergent sum $T_C(\rho) = \sum_i \langle e_i, \rho e_i\rangle M_{f_i}$, so the slice map takes values in $\{M_f : f \in C(L)\}$. Since λ has full support and f is continuous, $\|M_f\| = \sup|f|$; hence the map $f \mapsto M_f$ is an isometric $*$-isomorphism from $C(L)$ onto the image. The unital C*-subalgebra of $C(L)$ generated by the functions f_i is $C(L)$, by the Stone–Weierstrass theorem, because they separate points. It follows that $\mathfrak{A}_C = \{M_f : f \in C(L)\}$, and, by the Riesz representation theorem, the states of $\mathfrak{A}_C$ are the measures $\nu \in P(L)$, acting by $M_f \mapsto \int f \, d\nu$. Finally, if $\sigma = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$ is a density operator on $L^2(L,\lambda)$, with weights $\lambda_j \geq 0$ of sum 1 and unit vectors ψ_j, then $\operatorname{Tr}(\sigma M_f) = \int f \left(\sum_j \lambda_j |\psi_j|^2\right) d\lambda$, so the measure μ_σ has density $\sum_j \lambda_j |\psi_j|^2$ with respect to λ; conversely, a probability measure with density p with respect to λ is μ_σ for the pure state given by the unit vector $\sqrt{p}$. □

Transfer from points to measures. For a compact metrizable space L, a continuous function $g \colon L \to [0,1]$ determines a response $\mathfrak{r}_g \colon P(L) \to [0,1]$ by $\mathfrak{r}_g(\nu) = \int_L g \, d\nu$. To transfer the minimal families from functions on L to responses on the probe space, we must preserve their pointwise closures under this map. Lemma P.48 below does so for countable families whose pointwise closures consist of Baire-1 functions; the resulting sequential properties allow us to use dominated convergence and recover each limiting response from the values of the response on the point masses.

Lemma P.48. *Let L be a compact metrizable space and $(f_t)_{t\in T}$ a countable family of continuous functions $L \to [0,1]$ that is relatively compact in the space $B_1(L)$ of Baire-1 functions on L. Then the families $\{\mathfrak{r}_{f_t} : t \in T\}$, of functions on $P(L)$, and $\{f_t : t \in T\}$, of functions on L, are equivalent in the sense of Definition P.14, and the family $\{\mathfrak{r}_{f_t} : t \in T\}$ is NIP.*

Proof. Write $\mathfrak{r}_t = \mathfrak{r}_{f_t}$ for $t \in T$. Every sequence (f_{t_n}) has a subsequence that converges pointwise on L. Indeed, otherwise, by Fact P.9, some subsequence would have a closure homeomorphic to $\beta\mathbb{N}$, which is not Fréchet–Urysohn (a space is Fréchet–Urysohn if every point in the closure of a subset is the limit of a sequence from that subset), whereas the closure of $\{f_t\}$ is a Rosenthal compactum and therefore Fréchet–Urysohn (see [**BFT78**, Theorem 3F]). By the dominated convergence theorem, the corresponding subsequence of $(\mathfrak{r}_{t_n})$ converges pointwise on $P(L)$.

The family $\{\mathfrak{r}_t\}$ is thus NIP. Indeed, if a sequence $(\mathfrak{r}_{t_n})$ and reals $a < b$ witnessed the independence property, then, since $P(L)$ is compact and the functions $\mathfrak{r}_t$ are continuous, for each $E \subseteq \mathbb{N}$ there would be ν with $\mathfrak{r}_{t_n}(\nu) \le a$ for $n \in E$ and $\mathfrak{r}_{t_n}(\nu) \ge b$ for $n \notin E$ (see the proof of Theorem P.11); taking for E every other index of a pointwise convergent subsequence would yield a ν at which that subsequence does not converge, contrary to the previous paragraph. By Theorem P.11, the closure of $\{\mathfrak{r}_t\}$ is contained in $B_1(P(L))$; the closure is a Rosenthal compactum, hence Fréchet–Urysohn.

Let Φ send a function G on $P(L)$ to the function $x \mapsto G(\delta_x)$ on L. The map Φ is continuous for the pointwise topologies, and $\Phi(\mathfrak{r}_t) = f_t$. The map Φ is injective on the closure of $\{\mathfrak{r}_t\}$: if G lies in that closure, then, the closure being Fréchet–Urysohn, the function G is the pointwise limit of a sequence $(\mathfrak{r}_{t_n})$; the sequence (f_{t_n}) then converges pointwise to $\Phi(G)$, which is therefore a bounded Borel function, and, by dominated convergence, the value $G(\nu)$ equals $\lim_n \int f_{t_n}\, d\nu = \int \Phi(G)\, d\nu$ for every ν. Thus G is determined by $\Phi(G)$. As a continuous injection on a compact space, the map Φ is a homeomorphism of the closure of $\{\mathfrak{r}_t\}$ onto the image; this image is compact, contains $\{f_t\}$ and lies in the closure of $\{f_t\}$, so it is that closure. □

We also need to discard points that carry no information about limits.

Lemma P.49. *Let L be a set, $L_0 \subseteq L$, and $(g_t)_{t\in T}$ a countable pointwise bounded family of real functions on L whose pointwise closure K is Fréchet–Urysohn. Suppose that*

(i) *the restrictions $g_t|_{L_0}$ are pairwise distinct, and none of them lies in the closure of the others;*
(ii) *if two limits of sequences $(g_{t_k})_{k\in\mathbb{N}}$ with pairwise distinct indices agree on L_0, then they are equal.*

Then the families (g_t) and $(g_t|_{L_0})$ are equivalent.

Proof. Restriction $\pi\colon K \to \mathbb{R}^{L_0}$ is continuous, and the image of π is the closure of $(g_t|_{L_0})$. Let $f \in K$ be different from every g_t. Since K is Fréchet–Urysohn, the function f is the limit of a sequence (g_{t_k}); since K is Hausdorff and f is not a g_t, no index occurs infinitely often, and, after passage to a subsequence, the indices are pairwise distinct. By (i), the map π is injective on the family; moreover, $\pi(g_t) \ne \pi(f)$ for every such limit f and every t, because $\pi(f)$ lies in the closure of $\{g_s|_{L_0} : s \ne t\}$, whereas $g_t|_{L_0}$ does not. By (ii), the map π is also injective on the limits. Hence π is a continuous injection of the compact space K, and therefore a homeomorphism onto the image. □

Three settings. The realizations use three physical settings.

Photon numbers. This is the discrete comparison of section P.2, in which Alice's and Bob's values lie in $\Xi = \{x_l : l \in \mathbb{N}\} \cup \{1\}$, where $x_l = 1 - \frac{1}{l+1}$. For Alice's number state $|l\rangle$, the state with exactly l photons, the function $g_l = g_{|l\rangle}$, whose value at a point y is the probability that Alice's value lies below y, is the indicator of $\{x_j : j > l\} \cup \{1\}$; the probe space is the space $P(\Xi)$ of Borel probability measures on Ξ, and $\mathfrak{r}_{|l\rangle}(\nu) = \int g_l \, d\nu$ (see Proposition P.39).

Positions in a box. This is the continuous comparison of section P.2: the probe space is $P([0,1])$, and a state ρ of Alice's particle with distribution μ_ρ has response $\nu \mapsto \int g_\rho \, d\nu$, where $g_\rho(y) = \mu_\rho([0,y))$. We place a Cantor set in the box. Let $\psi(\sigma) = \frac{3}{4} - \sum_n \sigma(n) 3^{-n-1}$ for $\sigma \in \{0,1\}^{\mathbb{N}}$; this is an order-reversing homeomorphism of the Cantor space $\{0,1\}^{\mathbb{N}}$, with the lexicographic order $\prec$, onto a Cantor set $C_0 \subseteq [\frac{1}{4}, \frac{3}{4}]$. For a finite sequence $s \in \{0,1\}^{<\mathbb{N}}$, let I_s be the smallest interval containing the image $\psi(V_s)$ of the clopen set V_s of infinite extensions of s; the interval has length $\ell_s = \frac{1}{2} 3^{-|s|}$, left endpoint $\psi(s^\frown 1^\infty)$ and right endpoint $\psi(s^\frown 0^\infty)$, the images of the extensions of s by ones and by zeros. If ζ is the immediate successor of σ in $\prec$, then the interval $(\psi(\zeta), \psi(\sigma))$ is a component of $[0,1] \setminus C_0$; the other components are $[0, \frac{1}{4})$ and $(\frac{3}{4}, 1]$. The components adjacent to I_s have length at least $\ell_s/2$: each of them is either $[0, \frac{1}{4})$ or $(\frac{3}{4}, 1]$, of length $\frac{1}{4} \geq \ell_s/2$, or the gap between $I_{u^\frown 0}$ and $I_{u^\frown 1}$ for some proper initial segment u of s, of length $\ell_u/3 \geq \ell_s$.

Registers. Alice's system is a chain of spins of variable finite length, with Hilbert space $\ell^2(\{0,1\}^{<\mathbb{N}}) \cong \bigoplus_{n \in \mathbb{N}} (\mathbb{C}^2)^{\otimes n}$. For $t \in \{0,1\}^{<\mathbb{N}}$, the basis ket $|t\rangle = \bigotimes_{i<|t|} |t(i)\rangle$ has spin up at site i exactly when $t(i) = 0$, with $|0\rangle$ denoting up and $|1\rangle$ down. Bob probes the chain in two ways. In *scanning*, Bob's probe sits at a site $i \in \mathbb{N}$ and asks whether Alice's spin at that site is down, the outcome selected by $|1\rangle\langle 1|$: the test is $a_t(i) = t(i)$ for $i < |t|$ and $a_t(i) = 0$ otherwise. In *template matching*, Bob holds a reference configuration u, finite or infinite, and asks whether Alice's chain coincides with the beginning of his reference: the test is $b_t(u) = 1$ if $t \sqsubseteq u$ and $b_t(u) = 0$ otherwise. We extend a_t to the one-point compactification $L_a = \mathbb{N} \cup \{\infty\}$ by $a_t(\infty) = 0$, and we let L_b be the space of finite and infinite sequences of zeros and ones, with the ultrametric $d(u,u') = 2^{-|u \wedge u'|}$ for $u \neq u'$, where $u \wedge u'$ is the longest common initial segment; the space is compact, the finite sequences are isolated and dense, and the functions b_t are continuous, since $\{u : t \sqsubseteq u\}$ is clopen. The *scanning*, *template* and *combined* observables are the label-controlled observables with labels in $\{0,1\}^{<\mathbb{N}}$, on L_a, L_b and their disjoint sum $L_a \oplus L_b$, with the functions a_t, b_t and (a_t, b_t), the last equal to a_t on L_a and to b_t on L_b, and with measures concentrated on $\mathbb{N}$, on the finite sequences, and on both, respectively. These functions separate the points of L_a, of L_b and of $L_a \oplus L_b$; for instance, the functions of the empty label, $a_\emptyset = 0$ and $b_\emptyset = 1$, separate L_a from L_b. Hence Lemma P.47 applies to all three observables.

The realization. Together, the three settings realize all seven minimal families. In the next theorem, $(s_t)_{t \in \{0,1\}^{<\mathbb{N}}}$ is the canonical enumeration of the regular dyadic subtree Θ fixed in section P.2, and the chains $|t\rangle$, the Cantor set C_0, the homeomorphism ψ and the intervals I_s, of lengths ℓ_s, are those of the three settings.

Theorem P.50. *Each minimal family is realized by a tree of actual preparations of one of the observables of the three settings: for each $i = 1, \ldots, 7$, the observable and the pure states $(\rho_t)_{t \in \{0,1\}^{<\mathbb{N}}}$ of Alice's system listed below, given by pairwise orthogonal vectors, have responses $\{\mathfrak{r}_{\rho_t}\}$ equivalent to D_i.*

D_1: the comparison of photon numbers, with Alice's number states in any order;
D_2: scanning, with Alice's chains $|s_t\rangle$, whose scanning functions are $s_t^\frown 0^\infty$, for $s_t \in \Theta$;
D_3: the comparison of positions, with Alice's particle uniformly distributed over the gap of C_0 immediately to the right of $\psi(s_t^\frown 0^\infty)$;

D_4: *the comparison of positions, with Alice's particle uniformly distributed over the gap of C_0 immediately to the left of $\psi(s_t^\frown 1^\infty)$;*
D_5: *template matching, with Alice's chains $|t\rangle$;*
D_6: *the combined observable, with Alice's chains $|t\rangle$;*
D_7: *the comparison of positions, with Alice's particle distributed uniformly, half over $[\min I_{s_t} - \frac{\ell}{2}, \min I_{s_t} - \frac{\ell}{4}]$ and half over $[\max I_{s_t} + \frac{\ell}{4}, \max I_{s_t} + \frac{\ell}{2}]$, where $\ell = \ell_{s_t}$.*

Moreover, the causal order of spacetime realizes D_5 in the sense of Theorem P.42.

PROOF. In each case the probe space is $P(L)$ for a compact metrizable space L, the responses are $\mathfrak{r}_{\rho_t}(\nu) = \int g_t \, d\nu$ with g_t continuous, and the closure of $\{g_t\}$ consists of Baire-1 functions; by Lemma P.48, it suffices to show that $\{g_t\}$ is equivalent to D_i. We use the notation of section P.2: the function v_t is the characteristic function of the set V_t of infinite sequences that extend t, and the functions x_σ^+ and x_σ^- are the characteristic functions of $\{\zeta \in \{0,1\}^{\mathbb{N}} : \sigma \preceq \zeta\}$ and $\{\zeta \in \{0,1\}^{\mathbb{N}} : \sigma \prec \zeta\}$, for the lexicographic order $\prec$.

D_1. Index the functions g_l, $l \in \mathbb{N}$, by $\{0,1\}^{<\mathbb{N}}$ through any bijection, and write g_t for the function of index t. Every sequence of pairwise distinct members converges to the indicator of $\{1\}$, which is not a member, and each member is isolated. Hence the closure is a countable compact space with a single non-isolated point, and so is the closure $D_1 \cup \{0\}$ of D_1; the bijection $g_t \mapsto \frac{1}{|t|+1} v_t$, extended by sending the indicator of $\{1\}$ to 0, is a homeomorphism. The space Ξ is countable, so every function on it is Baire-1, every subset of a countable space being an F_σ.

D_2. Here $g_t = a_{s_t}$ is $s_t^\frown 0^\infty$ on $\mathbb{N}$ and 0 at ∞. This is the family D_2, extended continuously to $\mathbb{N} \cup \{\infty\}$; every function in the closure of this family vanishes at ∞, so the extension does not change the closure, and every function on the countable space L_a is Baire-1.

D_3 and D_4. Let $\sigma_t = s_t^\frown 0^\infty$. If $\sigma_t \neq 0^\infty$, the point σ_t has an immediate predecessor σ_t^-, and $G_t = (\psi(\sigma_t), \psi(\sigma_t^-))$ is a gap; if $\sigma_t = 0^\infty$, let $G_t = (\frac{3}{4}, 1]$. Let ρ_t be uniform on G_t. Then $g_t = 0$ on $[0, \psi(\sigma_t)]$, $g_t = 1$ to the right of G_t, and g_t is linear on G_t. On C_0, the function g_t satisfies $g_t(\psi(\zeta)) = 0$ exactly when $\zeta \succeq \sigma_t$; that is, $g_t \circ \psi = 1 - x_{\sigma_t}^+$. The sequences σ_t are pairwise distinct, and so are the gaps G_t. We apply Lemma P.49 with $L_0 = C_0$. The closure of $\{g_t\}$ consists of nondecreasing functions; the closure is therefore a Rosenthal compactum, and hence Fréchet–Urysohn. Condition (i) holds because D_3 is discrete in the closure of D_3. For (ii), let J be a component of $[0,1] \setminus C_0$. For pairwise distinct t_k, eventually $G_{t_k} \neq J$, and then the function g_{t_k} is constant on $\overline{J}$; every such limit f is thus constant on $\overline{J}$, and so is determined on J by the value of f at an endpoint of J in C_0. Consequently, the family $\{g_t\}$ is equivalent to $\{g_t|_{C_0}\}$, which is carried to D_3 by the homeomorphism ψ followed by the map $f \mapsto 1 - f$, a homeomorphism for the pointwise topology. The case of D_4 is symmetric: with $\sigma_t = s_t^\frown 1^\infty$ and the immediate successor σ_t^+, the gap is $(\psi(\sigma_t^+), \psi(\sigma_t))$, or $[0, \frac{1}{4})$ if $\sigma_t = 1^\infty$, and $g_t \circ \psi = 1 - x_{\sigma_t}^-$.

D_5. Here $g_t = b_t$ on L_b, and the restriction of g_t to the infinite sequences is v_t. Let K be the pointwise closure of $\{b_t\}$. For a finite sequence u, only finitely many functions b_t are nonzero at u; hence any member of K that is nonzero at u belongs to this finite set of functions. Every member of $K \setminus \{b_t : t \in \{0,1\}^{<\mathbb{N}}\}$ therefore vanishes on the finite sequences. The restriction to the closed set $\{0,1\}^{\mathbb{N}}$ of every member of K lies in $\hat{A}(2^{\mathbb{N}}) \subseteq B_1(\{0,1\}^{\mathbb{N}})$. The inverse image of an open set under such a member is therefore an F_σ subset of $\{0,1\}^{\mathbb{N}}$, together with at most countably many isolated points of L_b, and is thus F_σ in L_b. By Fact P.6(2), each member of K is Baire-1 on L_b. Thus K is a Rosenthal compactum and is Fréchet–Urysohn. We may now apply Lemma P.49 with $L_0 = \{0,1\}^{\mathbb{N}}$: condition (i) holds because D_5 is discrete

in the closure of D_5, and condition (ii) holds because limits of sequences with pairwise distinct indices vanish at every finite sequence. The family $\{b_t\}$ is consequently equivalent to D_5.

D_6. Here $g_t = (a_t, b_t)$ on $L_a \oplus L_b$. Every function on the countable compact metrizable space L_a is Baire-1, while every member of the pointwise closure of $\{b_t\}$ is Baire-1 on L_b, as shown for D_5. The pointwise closure of $\{g_t\}$ therefore consists of Baire-1 functions on the disjoint sum and is a Rosenthal compactum, hence Fréchet–Urysohn. On $L_0 = \mathbb{N} \oplus \{0,1\}^{\mathbb{N}}$, the function g_t is $(t^\frown 0^\infty, v_t)$, the corresponding member of D_6 with the two summands interchanged. These restrictions are discrete in their closure, and every limit of a sequence with pairwise distinct indices vanishes at ∞ and at each finite sequence in L_b. Thus both conditions of Lemma P.49 hold, and the family $\{g_t\}$ is equivalent to D_6.

D_7. Let $\ell = \ell_{s_t}$, and let ρ_t be the state whose distribution is half uniform on $A_t = [\min I_{s_t} - \frac{\ell}{2}, \min I_{s_t} - \frac{\ell}{4}]$ and half uniform on $B_t = [\max I_{s_t} + \frac{\ell}{4}, \max I_{s_t} + \frac{\ell}{2}]$; these intervals lie in the components adjacent to I_{s_t}. On C_0, the function g_t equals 0 to the left of I_{s_t}, $\frac{1}{2}$ on I_{s_t} and 1 to the right; that is, $g_t \circ \psi = 1 - x^+_{\sigma_t} + \frac{1}{2} v_{s_t}$ with $\sigma_t = s_t^\frown 0^\infty$, since $\psi(\sigma_t) = \max I_{s_t}$. Each member of the closure of D_7 has the form (v, x), for functions v and x on $\{0,1\}^{\mathbb{N}}$ with $v \le x$ that take only the values 0 and 1; hence the map $(v, x) \mapsto (1 - x + \frac{1}{2} v) \circ \psi^{-1}$ is a continuous injection, and the map carries D_7 to $\{g_t|_{C_0}\}$; as D_7 is discrete in the closure of D_7, condition (i) of Lemma P.49 holds. For (ii), note that A_t lies in the component whose right endpoint is $\psi(s_t^\frown 1^\infty)$, and B_t in the component whose left endpoint is $\psi(s_t^\frown 0^\infty)$. These sequences are pairwise distinct for $s_t \in \Theta$, so each component of $[0,1] \setminus C_0$ contains at most one interval A_t and at most one interval B_t. For pairwise distinct t_k, eventually neither A_{t_k} nor B_{t_k} lies in a given component J, and then the function g_{t_k} is constant on $\overline{J}$. Every such limit is therefore constant on $\overline{J}$ and is determined by the values of the limit on C_0.

Finally, the photon and register preparations are given by distinct basis vectors. In the box, take the unit vectors $\sqrt{d\mu_{\rho_t}/dy}$. Within each of the D_3 and D_4 realizations, the supports lie in distinct Cantor gaps. For D_7, each gap contains at most one interval A_t and at most one interval B_u, by the distinctness of the endpoint sequences fixed in section P.2. If an internal gap contains both, the length of the gap is at least each of ℓ_{s_t} and ℓ_{s_u}; the interval A_t lies in the right half of the gap and B_u in the left half, so they meet in at most the midpoint. Each exterior gap contains at most one support interval. Within each realization, the supports of distinct preparations consequently meet in at most finitely many points, whence the vectors are pairwise orthogonal in $L^2[0,1]$. □

The realizations display what distinguishes the families.

- D_1, *escape to infinity*: an oscillator excited ever higher has a single idealized limit, and each actual preparation is isolated from it.
- D_2, *infinite configurations*: the idealized limits of the chains are the infinite spin configurations, a Cantor set of them, and countably many site probes tell them all apart, so the closure is metrizable; but any two configurations differ by 1 at some site, so the closure is not separable in the supremum norm.
- D_3 and D_4, *one-sided localization*: a particle that approaches a point from only one side produces one half of a split pair, and the split Cantor set appears; the closure is hereditarily separable but not metrizable.
- D_5, *point-supported limits*: light cones that shrink to a point on Bob's slice, or templates following longer and longer initial segments of one infinite configuration, yield a point-supported limiting profile; the corresponding idealized response evaluates the mass that a probe assigns to the point, and the closure of the selected responses is not first countable.

- D_6, *twins over a metrizable base*: each infinite configuration has two idealized versions, one that every template misses and one that the template of the configuration itself detects; the scanning probes alone cannot tell the twins apart, and, as seen by these probes, the limits form a Cantor set.
- D_7, *twins over a split base*: a particle that approaches a point from both sides at once produces the midpoint of a split pair, as the symmetric state does in the Ising model (see section P.2 below); the twins of the closure sit over the split Cantor set.

The scanning and template observables are NIP (see Proposition P.51 and Corollary P.52 below), which places them in the classification of Theorem P.32; the combined observable is NIP by Proposition P.51 as well, since the functions of the observable separate points and every sequence of them has a subsequence that converges pointwise on each of the two summands (see the proof of Corollary P.52). The proof of Proposition P.51 uses a theorem of Stegall. Recall that a subset of a Banach space is *weakly precompact* if it is bounded and every sequence in it has a weakly Cauchy subsequence, one whose values under each continuous linear functional form a Cauchy sequence, and that the *absolutely convex hull* of a set consists of the finite combinations $\sum_k a_k x_k$ of elements of the set with $\sum_k |a_k| \le 1$. Stegall's theorem states that the absolutely convex hull of a weakly precompact set is weakly precompact (see Rodríguez [**Rod22**] for a quantitative version). We say that an observable is *NIP for Bob's responses* if the observable obtained from it by exchanging the two tensor factors is NIP for Alice's responses.

Proposition P.51. *Let C be a label-controlled observable whose functions f_i separate the points of L and have the property that every sequence of them has a subsequence converging pointwise on L. Then the observable C is NIP for Alice's responses.*

Proof. For a density operator ρ, let $p_i = \langle e_i, \rho e_i \rangle$, a probability distribution on I, and $g_\rho = \sum_i p_i f_i$, a norm-convergent sum in $C(L)$; then $\mathfrak{r}_\rho(\nu) = \int g_\rho \, d\nu$. By dominated convergence, a sequence of functions f_i that converges pointwise on L is weakly Cauchy in $C(L)$; hence the set $\{f_i\}$ is weakly precompact. By Stegall's theorem, so is the absolutely convex hull of the set, and so is the norm closure of that hull: given y_n in the closure, choose x_n in the hull with $\|x_n - y_n\| < 1/n$ and a weakly Cauchy subsequence of (x_n); the corresponding subsequence of (y_n) is weakly Cauchy. The partial sums of g_ρ lie in the hull, and every function g_ρ therefore lies in this closure. Consequently, every sequence (g_{ρ_n}) has a weakly Cauchy subsequence, and, since every $\nu \in P(L)$ is a continuous linear functional on $C(L)$, the corresponding subsequence of $(\mathfrak{r}_{\rho_n})$ converges pointwise on $P(L) = Y_C$. As in the proof of Lemma P.48, an independent sequence has no such subsequence. □

Corollary P.52.

(a) *The scanning observable is unstable and NIP$^{\mathrm{met}}$, but not uniformly NIP.*
(b) *The template observable is NIP but not NIP$^{\mathrm{fc}}$.*
(c) *The scanning observable is not NIP for Bob's responses. Hence NIP for Alice's responses does not imply NIP for Bob's responses.*

Proof. (a) Every sequence of functions a_t on the countable set L_a has a pointwise convergent subsequence, by a diagonal argument, so the observable is NIP by Proposition P.51. For a trace-norm dense sequence $(\rho_n)_{n\in\mathbb{N}}$, the closure of the responses is that of $\{\mathfrak{r}_{\rho_n}\}$ (see Lemma P.26). Every function on the countable compact metrizable space L_a is Baire-1, so Lemma P.48 applies and identifies this closure with the closure of $\{g_{\rho_n}\}$ in $\mathbb{R}^{L_a}$; since L_a is countable, this space is metrizable. For each positive integer m, the chains $t_1, \dots, t_m$ of length 2^m with $t_j(i)$ the j-th binary digit of i are shattered, with values 0 and 1, by the site

probes $i < 2^m$; hence the observable is not uniformly NIP. Dually, on the first m site probes, the chains of length m realize every binary pattern; the fat-shattering dimension of the response family is therefore infinite at every positive gap at most 1, with the gap convention of Remark P.25. Finally, the observable is unstable: the functions a_{1^k} of the chains 1^k of k down spins, for $k \in \mathbb{N}$, converge pointwise to the indicator of $\mathbb{N}$, so the idealized response $\nu \mapsto \nu(\mathbb{N})$, which is not continuous on $P(L_a)$, is a limit of responses (see Theorem 11.1(a)).

(b) Let $(t_k)_{k\in\mathbb{N}}$ be a sequence of labels in $\{0,1\}^{<\mathbb{N}}$. By compactness of L_b, there is a strictly increasing map $j\colon \mathbb{N} \to \mathbb{N}$ such that $(t_{j(k)})_{k\in\mathbb{N}}$ converges in L_b. If the limit of this subsequence is finite, the subsequence is eventually constant, so the corresponding functions converge pointwise. If the limit is an infinite sequence $\sigma \in \{0,1\}^{\mathbb{N}}$, then the values satisfy $b_{t_{j(k)}}(u) \to 0$ for each finite sequence u, and $b_{t_{j(k)}}(\zeta) \to 0$ for each infinite sequence $\zeta \neq \sigma$: eventually $t_{j(k)}$ agrees with σ beyond the first place where σ and ζ differ. Passing to a further subsequence on which $b_{t_{j(k)}}(\sigma)$ is constant gives pointwise convergence on all of L_b. Hence the observable is NIP by Proposition P.51. By Theorem P.50, the closure of the responses of the observable contains a closed copy of $\hat{A}(2^{\mathbb{N}})$, which is not first countable.

(c) For $i \in \mathbb{N}$, consider the state of Bob's system in which his probe sits at the site i; these states form an independent sequence: for disjoint finite sets $E, F \subseteq \mathbb{N}$, choose a positive integer q larger than every member of $E \cup F$ and let Alice's chain t of length q have $t(i) = 1$ exactly for $i \in F$. Then $a_t(i) = 1$ for $i \in F$ and $a_t(i) = 0$ for $i \in E$. Alice's actual states restrict to points of the probe space of the exchanged observable (see Lemma P.22(c)), so the family of Bob's responses has the independence property. This is the relation $i \in u$ of Remark P.13, realized by a physical observable. □

The scanning observable shows the difference between NIP and uniform NIP (see Remark P.25): the idealized responses form a metrizable compactum, yet the response family has infinite fat-shattering dimension at a fixed positive gap. In distribution-free learning from independent samples, the aim is to predict an unknown response with small mean absolute error under the probe distribution, with a prescribed probability of success. No sample bound depending only on the required accuracy and probability of success can achieve this for the scanning family uniformly over target responses and probe distributions [**BLW96, ABDCBH97**]. The obstruction is already visible on the first m sites, where chains of length m realize every binary pattern.

Spontaneous symmetry breaking is a split pair. We now show that the levels of section P.2 have a physical meaning beyond bipartite observables. Several notions of physics rest on an order of limits of a family of responses: a preparation, such as a finite sample, has a response, such as its magnetization as a function of an external field, and the notion depends on whether the sample becomes infinite before or after the field is switched off.

The common setting is a pointwise bounded family $\mathcal{R}$ of continuous functions on a compact metrizable probe space Y; the idealized responses of the family are the members of the pointwise closure $\overline{\mathcal{R}}$. We call the family *stable* when, for every sequence $(r_k)_{k\in\mathbb{N}}$ in $\mathcal{R}$ and every sequence $(y_l)_{l\in\mathbb{N}}$ in Y, the iterated limits $\lim_l \lim_k r_k(y_l)$ and $\lim_k \lim_l r_k(y_l)$ agree whenever both exist. For a uniformly bounded family, as for the magnetizations below, Grothendieck's double limit exchange condition [**Gro52**] identifies stability with continuity of every idealized response; countably many probes then separate the closure, which is metrizable by Proposition P.18. The NIP characterization of Theorem P.11 applies to pointwise bounded families, and Definition P.17 assigns their finer levels when the closure is separable. For a fixed observable, convexity follows from allowing all mixed preparations; for the parameterized families

considered here, we compute the closures directly from the stated properties of the physical models.

The standard example of an order of limits is spontaneous symmetry breaking: for a ferromagnet below its critical temperature, letting the external field vanish after the volume becomes infinite yields a nonzero magnetization, whereas the other order yields zero [**FV17**]. This is an instability in our sense, absent above the critical temperature (see Theorem P.54 below). With an offset field that moves the transition point, the field at which the magnetization changes sign, the family of responses below the critical temperature is moreover $\mathrm{NIP}^{\mathrm{fc}}$ but not $\mathrm{NIP}^{\mathrm{hs}}$: the idealized responses of the two pure phases, the equilibrium states magnetized in opposite directions, form a split pair, two functions that differ only at the transition point, and the idealized response of the symmetric state, the equal mixture of the two phases, is their midpoint, as in the example before Theorem P.29.

The Ising model. We recall the facts we need; all of them can be found in the book of Friedli and Velenik [**FV17**]. Fix an integer dimension $d \geq 2$, a coupling energy $J > 0$ and an inverse temperature $\beta = 1/(k_B T) > 0$, where k_B is the Boltzmann constant and T the temperature. The external field h is a real number in units of energy, so βJ and βh are dimensionless. For each positive integer half-width N, let $\Lambda_N = \{-N, \dots, N\}^d$ with periodic boundary conditions, and define $Z_N \colon \mathbb{R} \to (0, \infty)$ and $p_N \colon \mathbb{R} \to \mathbb{R}$ by

$$Z_N(h) = \sum_{s \in \{-1,1\}^{\Lambda_N}} \exp\Big(\beta J \sum_{i \sim j} s_i s_j + \beta h \sum_i s_i\Big), \qquad p_N(h) = \frac{1}{|\Lambda_N|} \log Z_N(h),$$

where $i \sim j$ ranges over the unordered nearest-neighbor pairs of the torus, counted once, and the sum over i ranges over Λ_N; the corresponding energy of a configuration s is $E_{N,h}(s) = -J \sum_{i \sim j} s_i s_j - h \sum_i s_i$. The Gibbs state is the probability measure on configurations that gives s a weight proportional to $e^{-\beta E_{N,h}(s)}$; the normalizing constant $Z_N(h)$ is the *partition function*, and the function p_N is the *pressure* of the finite system, that is, $-\beta$ times the free energy per site. The magnetization per site in the Gibbs state is the function $m_N \colon \mathbb{R} \to [-1, 1]$ given by $m_N(h) = p_N'(h)/\beta$.

(F1) The function p_N is smooth, convex and even. Hence the function m_N is continuous, nondecreasing and odd, with $m_N(0) = 0$. The bound $|m_N| \leq 1$ holds because $m_N(h)$ is the mean, in the Gibbs state, of the average spin $|\Lambda_N|^{-1} \sum_i s_i$, which lies in $[-1, 1]$.

(F2) The limit $p(h) = \lim_N p_N(h)$, the pressure, exists for every h; the pressure is convex and even, and is differentiable at every $h \neq 0$, by the Lee–Yang theorem [**LY52**]. Since p_N and p are convex, $m_N(h) \to p'(h)/\beta$ at every $h \neq 0$, by Griffiths' lemma on derivatives of convex functions. Indeed, convexity bounds $p_N'(h)$ between the secant slopes over $[h - \eta, h]$ and $[h, h + \eta]$ for $\eta > 0$; first let $N \to \infty$, then $\eta \to 0$ at a differentiability point of p.

(F3) The spontaneous magnetization $m^*(\beta) = p'(0^+)/\beta$ is the right derivative of p at 0 divided by β, and $p'(0^-)/\beta = -m^*(\beta)$. With J fixed, the critical inverse temperature $\beta_c = \inf\{\beta : m^*(\beta) > 0\}$ is positive, and finite by Peierls's argument [**Pei36**, **Gri64**]; since $m^*(\beta)$ is the magnetization of the plus state (the infinite-volume limit of the Gibbs states with all boundary spins fixed up), the spontaneous magnetization is nondecreasing in β, by the Griffiths–Kelly–Sherman inequalities, and $m^*(\beta) > 0$ for $\beta > \beta_c$ (see [**FV17**, Chapter 3]).

As p'/β is nondecreasing and odd, (F3) implies that $p'(t)/\beta \geq m^*$ for $t > 0$ and $p'(t)/\beta \leq -m^*$ for $t < 0$.

Preparations and probes. A preparation consists of a finite system Λ_N and an offset field $a \in [-1/\beta, 1/\beta]$, that is, a uniform field built into the sample; a probe is an applied field $h \in Y = [-2/\beta, 2/\beta]$, a value of an external parameter rather than a preparation of a second system. The system then feels the field $h - a$, and the response of the preparation (N, a) is the function $r_{N,a} \colon Y \to [-1, 1]$ given by

$$r_{N,a}(h) = m_N(h - a) \qquad (h \in Y).$$

The offset is a device that moves the transition point; without it, the family $\{m_N|_Y : N \geq 1\}$ has only one possible transition point, at $h = 0$. Let $\mathcal{R} = \{r_{N,a} : N \geq 1,\ a \in [-1/\beta, 1/\beta]\}$. Each $r_{N,a}$ is continuous and nondecreasing, and $r_{N,a'} \to r_{N,a}$ uniformly as $a' \to a$, because m_N is uniformly continuous on compact sets. The responses with βa rational are therefore dense in $\mathcal{R}$, and the closure $\overline{\mathcal{R}}$ is separable. Every function in $\overline{\mathcal{R}}$ is nondecreasing, and so of the first Baire class; thus $\overline{\mathcal{R}}$ is a Rosenthal compactum, and, by Theorem P.11, the family $\mathcal{R}$ is NIP.

We write $m^* = m^*(\beta)$, and for $a \in [-1/\beta, 1/\beta]$ and $c \in [-m^*, m^*]$ we let $g_{a,c} \colon Y \to \mathbb{R}$ be given by $g_{a,c}(h) = p'(h - a)/\beta$ for $h \neq a$ and $g_{a,c}(a) = c$. These functions describe the idealized responses.

Lemma P.53. *The following hold:*

(i) *every $g_{a,c}$ with $-1/\beta < a < 1/\beta$ belongs to $\overline{\mathcal{R}}$;*
(ii) *every function in $\overline{\mathcal{R}} \setminus \mathcal{R}$ equals $g_{a,c}$ for some $a \in [-1/\beta, 1/\beta]$ and $c \in [-m^*, m^*]$.*

Proof. (i) Let $-1/\beta < a < 1/\beta$. If $m^* > 0$, then, since $\overline{\mathcal{R}}$ is closed and $g_{a,c} \to g_{a,\pm m^*}$ pointwise as $c \to \pm m^*$, it suffices to take $|c| < m^*$, and by symmetry $0 \leq c < m^*$; if $m^* = 0$, then $c = 0$. If $c = 0$, then $r_{N,a} \to g_{a,0}$, by (F1) at $h = a$ and by (F2) elsewhere. If $0 < c < m^*$, then, for every $\eta > 0$, we have $m_N(\eta) \to p'(\eta)/\beta \geq m^* > c$, by (F2) and (F3). Since $m_N(0) = 0$, for all large N the intermediate value theorem yields $\varepsilon_N \in (0, 1/\beta)$ with $m_N(\varepsilon_N) = c$; for each $\eta \in (0, 1/\beta)$, eventually $m_N(\eta) > c$, and then $\varepsilon_N < \eta$, because m_N is nondecreasing. Hence $\varepsilon_N \to 0$. Let $a_N = a - \varepsilon_N$. Then $r_{N,a_N}(a) = m_N(\varepsilon_N) = c$. For $h \neq a$, let $t = h - a$ and fix η' with $0 < \eta' < |t|$; for large N, we have $t - \eta' \leq h - a_N \leq t + \eta'$, and hence, by monotonicity, $m_N(t - \eta') \leq r_{N,a_N}(h) \leq m_N(t + \eta')$. By (F2), the outer terms converge to $p'(t - \eta')/\beta$ and $p'(t + \eta')/\beta$, which tend to $p'(t)/\beta$ as $\eta' \to 0$, because p' is continuous at $t \neq 0$. Finally, as $a > -1/\beta$, the offsets a_N lie in $[-1/\beta, 1/\beta]$ for large N, and $r_{N,a_N} \to g_{a,c}$.

(ii) Let $g \in \overline{\mathcal{R}} \setminus \mathcal{R}$. The Rosenthal compactum $\overline{\mathcal{R}}$ is Fréchet–Urysohn [**BFT78**, Theorem 3F], so the function g is the limit of a sequence r_{N_k,a_k}. Passing to a subsequence, we may assume that $a_k \to a$ and that either N_k is constant or $N_k \to \infty$. In the first case, uniform continuity forces $g = r_{N,a} \in \mathcal{R}$, which is excluded. In the second case, the argument of the previous paragraph shows that $g(h) = p'(h - a)/\beta$ for $h \neq a$. Finally, given $\eta > 0$, we have $|a - a_k| < \eta$ for all large k, and then $|r_{N_k,a_k}(a)| = |m_{N_k}(a - a_k)| \leq m_{N_k}(\eta)$, because m_{N_k} is odd and nondecreasing; the right side tends to $p'(\eta)/\beta$, and $p'(\eta)/\beta \to m^*$ as $\eta \to 0$. Hence $|g(a)| \leq m^*$, and $g = g_{a,g(a)}$. □

The classification detects the phase transition. By Lemma P.53, the classification of the family $\mathcal{R}$ of responses with offsets reduces to the behavior of p' at 0.

Theorem P.54. *Let $d \geq 2$ and $\beta > 0$ be given, let $m^*(\beta)$ be the spontaneous magnetization and β_c the critical inverse temperature of (F3), and let $\mathcal{R}$ be the family of responses $r_{N,a}$ with offsets.*

(a) *If $m^*(\beta) = 0$, in particular if $\beta < \beta_c$, then every idealized response is continuous: the family $\mathcal{R}$ is stable, and $\overline{\mathcal{R}}$ is metrizable.*
(b) *If $m^*(\beta) > 0$, in particular if $\beta > \beta_c$, then the family $\mathcal{R}$ is unstable, and $\overline{\mathcal{R}}$ is first countable but not hereditarily separable: the family is NIP$^{\mathrm{fc}}$ but not NIP$^{\mathrm{hs}}$.*

The family $\{m_N|_Y : N \geq 1\}$ without offsets is stable exactly when $m^*(\beta) = 0$, and is NIP$^{\mathrm{met}}$ for every β.

Proof. (a) If $m^* = 0$, then the pressure p is differentiable everywhere, and the derivative p', being the derivative of a convex function, is continuous. By Lemma P.53(ii), every idealized response outside $\mathcal{R}$ has the form $g_{a,0}(h) = p'(h-a)/\beta$ and is therefore continuous; so are the responses in $\mathcal{R}$. By Grothendieck's theorem [**Gro52**], a uniformly bounded family of continuous functions on a compact space whose pointwise closure consists of continuous functions satisfies the double limit condition, so the family $\mathcal{R}$ is stable. The closure is metrizable by the argument of Proposition P.28: evaluation at the rational points of Y is a continuous injection on continuous functions.

(b) Let $m^* > 0$. For the family without offsets, take $h_l = 1/(\beta l)$ for $l \geq 1$ and a sequence $(N_k)_{k\in\mathbb{N}}$ of positive integers with $N_k \to \infty$. By (F2) and (F3), $\lim_l \lim_k m_{N_k}(h_l) = \lim_l p'(1/(\beta l))/\beta = m^*$, whereas, by continuity of each m_N and (F1), $\lim_k \lim_l m_{N_k}(h_l) = \lim_k m_{N_k}(0) = 0$. The family $\{m_N\}$, and with it the family $\mathcal{R}$, is therefore unstable.

Every function in $\overline{\mathcal{R}}$ is nondecreasing, and the argument in the proof of Theorem P.40(b) shows that the space of nondecreasing functions $Y \to [-1,1]$ is first countable; hence so is $\overline{\mathcal{R}}$. For $a \in (-1/\beta, 1/\beta)$, the neighborhood $\{g : |g(a)| < m^*/2\}$ of $g_{a,0}$ contains no $g_{b,0}$ with $b \neq a$, because $|g_{b,0}(a)| = |p'(a-b)/\beta| \geq m^*$. The set $\{g_{a,0} : -1/\beta < a < 1/\beta\}$ is thus an uncountable discrete subspace, and $\overline{\mathcal{R}}$ is not hereditarily separable.

The first part of the last statement follows from part (a), since $\{m_N\} = \{r_{N,0}\}$ is a subfamily of $\mathcal{R}$, and from the computation at the beginning of (b). For the second part, the sequence (m_N) converges pointwise, by (F1) and (F2), so the closure of the sequence is countable, and a countable compact Hausdorff space is metrizable. □

For an interior offset $a \in (-1/\beta, 1/\beta)$, the proof describes the limiting responses at the transition point $h = a$. At $h = a$, the idealized responses take every value c between $-m^*$ and m^*. Suppose first that $m^* > 0$. The extreme values $\pm m^*$ are the magnetizations of the two pure phases, obtained by letting the field reach the transition from one side after the volume has become infinite; they are the values at a of the idealized responses g_{a,m^*} and $g_{a,-m^*}$, which form a split pair. The value $c = (2\lambda - 1)m^*$ is the magnetization of the mixture of the two phases with weights λ and $1-\lambda$, for $0 \leq \lambda \leq 1$, and this value is reached through finite-size rounding of the transition, the smoothing of the jump of the magnetization in a finite sample [**BK90**], by tuning the offset as the volume grows. As a varies over $(-1/\beta, 1/\beta)$, the symmetric responses $g_{a,0}$ form the uncountable discrete set that destroys hereditary separability. They are the midpoints of the split pairs, as in the example before Theorem P.29. If $m^* = 0$, in particular above the critical temperature, the two extreme values coincide, and all idealized responses are continuous. In this sense the classification detects the phase transition: passing from $\beta < \beta_c$ to $\beta > \beta_c$ moves the family with offsets from the stable, metrizable class to the class of NIP$^{\mathrm{fc}}$ families that are not NIP$^{\mathrm{hs}}$, the class of the compactum of idealized responses of the comparison of positions in a box, a closed subspace of the Helly space. Without offsets, the family stays NIP$^{\mathrm{met}}$, and only stability registers the transition.

The quantum Ising model. The analogue of Theorem P.54 holds for a Hamiltonian whose terms do not commute, the Ising model in a transverse field (see Theorem P.58 below). The only ingredient of the classical proof that requires an argument for this model is the differentiability of the pressure away from zero field. It can be derived from the Griffiths–Hurst–Sherman inequality, which holds for this model through a space-time representation, a description of the quantum model by random configurations in one more dimension, that of imaginary time [**BG09**, Lemma 4.6]; we derive the differentiability instead from a Lee–Yang theorem, which yields more (see Propositions P.56 and P.57 below): the partition function has no zeros off the imaginary axis of the complex longitudinal field (the field along the axis of the couplings), and the pressure is real analytic away from zero field.

Fix a positive integer d, a coupling $J > 0$, a transverse field $\Gamma \geq 0$ and an inverse temperature $\beta > 0$; the coupling and the transverse field are energies, and the inverse temperature is the reciprocal of the thermal energy $k_B T$, so that βJ and $\beta\Gamma$ are pure numbers. For each positive integer half-width N, take the torus $\Lambda_N = \{-N, \dots, N\}^d$ with periodic boundary conditions and Hilbert space $(\mathbb{C}^2)^{\otimes\Lambda_N}$. The Pauli matrices $\sigma^x = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\sigma^z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ act at each site $i \in \Lambda_N$ as σ_i^x and σ_i^z, respectively, with the identity on the other tensor factors. Let $S_N = \sum_{i\in\Lambda_N} \sigma_i^z$. Define the Hamiltonian and the entire function $Z_N \colon \mathbb{C} \to \mathbb{C}$ by

$$H_N = -J \sum_{i\sim j} \sigma_i^z \sigma_j^z - \Gamma \sum_i \sigma_i^x, \qquad Z_N(h) = \operatorname{Tr} \exp\bigl(-\beta(H_N - hS_N)\bigr),$$

where $h \in \mathbb{C}$, the bond sum runs over unordered nearest-neighbor pairs, counted once, and the site sum runs over Λ_N; the longitudinal field h is an energy, as in the classical model. For real h, the partition function is positive. Define the pressure, $-\beta$ times the free energy per site, by $p_N \colon \mathbb{R} \to \mathbb{R}$, $p_N(h) = |\Lambda_N|^{-1} \log Z_N(h)$. Define the magnetization $m_N \colon \mathbb{R} \to [-1, 1]$ by $m_N = p_N'/\beta$. For $\Gamma = 0$ this is the classical model with the same J, β and h.

(Q1) The function p_N is smooth, convex and even, and $m_N(h)$ is the expectation of $S_N/|\Lambda_N|$ in the Gibbs state, the density operator $e^{-\beta(H_N - hS_N)}/Z_N(h)$; hence the function m_N is continuous, nondecreasing and odd, with $|m_N| \leq 1$, because $\|S_N\| = |\Lambda_N|$, and $m_N(0) = 0$. For self-adjoint matrices X and S and real u, write $A = X + uS$. The matrix derivative formula

$$\frac{d}{du} e^{X+uS} = \int_0^1 e^{(1-t)A} S e^{tA}\, dt$$

and cyclicity of the trace give $\frac{d}{du} \operatorname{Tr} e^{X+uS} = \operatorname{Tr}(S e^{X+uS})$. By the Gibbs variational formula, the function $u \mapsto \log \operatorname{Tr} e^{X+uS} = \sup_\rho \{\operatorname{Tr} \rho(X + uS) - \operatorname{Tr} \rho \log \rho\}$, where the supremum runs over density operators ρ and $0 \log 0 = 0$, is convex as a supremum of affine functions. Taking $X = -\beta H_N$, $S = S_N$ and $u = \beta h$ gives $p_N'(h) = \beta m_N(h)$. The unitary $U = \prod_i \sigma_i^x$ commutes with H_N and satisfies $U S_N U^* = -S_N$, so $Z_N(-h) = Z_N(h)$ for all complex h.

(Q2) The limit $p(h) = \lim_N p_N(h)$ exists for every real h; it is convex and even. For free boundary conditions, in which the bonds that wrap around the torus are omitted, the existence is standard [**Sim93**], and the periodic Hamiltonian differs in operator norm from the free one by at most J times the number of these bonds, which is at most $d(2N + 1)^{d-1}$, so the two pressures differ by at most $\beta J d/(2N + 1)$, because $|\log \operatorname{Tr} e^A - \log \operatorname{Tr} e^B| \leq \|A - B\|$ for self-adjoint matrices A and B.

The proof of Proposition P.56 below uses the Suzuki–Trotter mapping [**Suz76**] between quantum spin systems and classical Ising systems in one more dimension to show that the

partition function of the quantum model is a limit of positive multiples of partition functions of classical Ising models on a larger graph, in which each site of the torus is repeated M times, for a positive integer M that tends to infinity, one copy for each slice of "imaginary time". All couplings, between neighboring sites of the torus and between consecutive copies of a site, are ferromagnetic, so the classical Lee–Yang theorem applies to each of these classical models: none of their partition functions vanishes off the imaginary axis. A limit of holomorphic functions without zeros has no zeros either, unless it vanishes identically (see Fact P.55(b)), and so the quantum partition function inherits the zero-free half-planes. The analyticity of the pressure then follows, in Proposition P.57, from a uniform bound and two classical theorems of complex analysis. The two proofs use only the following standard facts.

Fact P.55.

(a) Lee–Yang theorem [**LY52**] *(see* [**FV17**, Chapter 3]*). Let G be a finite graph with vertex set V and edge set E, let $J_e \geq 0$ for $e \in E$, and let h_v be complex numbers with $\operatorname{Re} h_v > 0$ for $v \in V$. Then $\sum_{s\in\{-1,1\}^V} \exp\bigl(\sum_{e=\{u,w\}\in E} J_e s_u s_w + \sum_{v\in V} h_v s_v\bigr) \neq 0$.*

(b) Hurwitz's theorem [**Tit39**]. *If holomorphic functions without zeros on a connected open set $U \subseteq \mathbb{C}$ converge uniformly on compact subsets of U to f, then either f has no zeros in U or f vanishes identically.*

(c) Borel–Carathéodory inequality [**Tit39**]. *If f is holomorphic on a neighborhood of the closed disc $|z-z_0| \leq R$ and $\operatorname{Re} f \leq A$ there, then $|f(z)-f(z_0)| \leq \frac{2r}{R-r}\bigl(A-\operatorname{Re} f(z_0)\bigr)$ for $|z-z_0| \leq r < R$.*

(d) Vitali's theorem [**Tit39**]. *A sequence of holomorphic functions on a connected open set U that is uniformly bounded on each compact subset of U and converges at every point of a set with an accumulation point in U converges uniformly on compact subsets of U; by Weierstrass's theorem, so do the derivatives.*

Proposition P.56 (Lee–Yang theorem for the quantum Ising model). *For every N, the function Z_N is entire, $Z_N(h) \neq 0$ whenever $\operatorname{Re} h \neq 0$, and $|Z_N(h)| \leq Z_N(\operatorname{Re} h)$.*

Proof. Write $X = \beta J \sum_{i\sim j} \sigma_i^z \sigma_j^z + \beta h S_N$, which is diagonal in the basis $\{|s\rangle : s \in \{-1,1\}^{\Lambda_N}\}$ of eigenvectors of the operators σ_i^z, and $W = \beta\Gamma \sum_i \sigma_i^x$. The function $Z_N(h) = \operatorname{Tr} e^{X+W}$ is entire, because the exponential series converges locally uniformly in h.

Step 1: the Lie product formula. By the Lie product formula

$$e^{A+B} = \lim_{M\to\infty} \bigl(e^{A/M} e^{B/M}\bigr)^M,$$

which holds uniformly for matrices A and B in a bounded set, the functions $Z_{N,M}(h) = \operatorname{Tr}(e^{X/M} e^{W/M})^M$ converge to $Z_N(h)$ as $M \to \infty$, uniformly for h in compact subsets of $\mathbb{C}$.

Step 2: a classical ferromagnet. Suppose first that $\Gamma > 0$, and let $\varepsilon = \beta\Gamma/M$. Since $e^{\varepsilon\sigma^x} = \cosh\varepsilon\, I + \sinh\varepsilon\, \sigma^x$, the matrix element $\langle s|e^{\varepsilon\sigma^x}|s'\rangle$ of a single spin is $\cosh\varepsilon$ if $s = s'$ and $\sinh\varepsilon$ if $s \neq s'$, that is, $\langle s|e^{\varepsilon\sigma^x}|s'\rangle = c_\varepsilon e^{K_\varepsilon s s'}$ with $c_\varepsilon = (\sinh\varepsilon\cosh\varepsilon)^{1/2} > 0$ and $K_\varepsilon = \frac{1}{2}\log\coth\varepsilon > 0$. Insertion of the basis between the factors of $(e^{X/M}e^{W/M})^M$ shows that, for $M \geq 3$,

$$Z_{N,M}(h) = c_\varepsilon^{M|\Lambda_N|} \sum_{s^1,\dots,s^M} \exp \sum_{k=1}^{M} \Bigl(\frac{\beta J}{M} \sum_{i\sim j} s_i^k s_j^k + K_\varepsilon \sum_i s_i^k s_i^{k+1} + \frac{\beta h}{M} \sum_i s_i^k \Bigr),$$

where $s^1, \dots, s^M$ range over $\{-1,1\}^{\Lambda_N}$ and $s^{M+1} = s^1$. This is $c_\varepsilon^{M|\Lambda_N|}$ times the partition function of a classical Ising model on the finite graph $\Lambda_N \times \{1,\dots,M\}$, with nonnegative couplings and the same dimensionless field $\beta h/M$ at every site. By Fact P.55(a), the function

does not vanish when $\operatorname{Re} h > 0$. Moreover, by the triangle inequality applied to the sum, $|Z_{N,M}(h)| \le Z_{N,M}(\operatorname{Re} h)$. If $\Gamma = 0$, then the function Z_N itself is a classical Ising partition function, and the same two statements hold directly.

Step 3: Hurwitz's theorem. The functions $Z_{N,M}$ have no zeros in the connected open set $\{\operatorname{Re} h > 0\}$ and converge to Z_N uniformly on the compact subsets of this set, and Z_N is not identically zero there, since it is positive on the positive real axis. By Fact P.55(b), the function Z_N has no zeros in $\{\operatorname{Re} h > 0\}$, and by the evenness in (Q1) none in $\{\operatorname{Re} h < 0\}$. The inequality $|Z_N(h)| \le Z_N(\operatorname{Re} h)$ follows from Step 2 in the limit. □

Proposition P.57. *The pressure p is real analytic on $\mathbb{R} \setminus \{0\}$, and $m_N(h) \to p'(h)/\beta$ for every $h \neq 0$.*

Proof. By evenness it suffices to treat $h > 0$. The half-plane $\{\operatorname{Re} h > 0\}$ is simply connected and Z_N does not vanish there, so the function Z_N has a holomorphic logarithm on it, and we let f_N be $|\Lambda_N|^{-1}$ times the logarithm that is real on $(0,\infty)$; thus $f_N = p_N$ on $(0,\infty)$ and $\operatorname{Re} f_N(h) = |\Lambda_N|^{-1} \log |Z_N(h)| \le p_N(\operatorname{Re} h)$, by Proposition P.56.

We bound p_N on the real line uniformly in N. The torus has $d|\Lambda_N|$ bonds, whence $\|\beta H_N\| \le \beta(dJ + \Gamma)|\Lambda_N|$. Since the trace of the exponential of a self-adjoint matrix of size $2^{|\Lambda_N|}$ and norm at most c lies between $2^{|\Lambda_N|}e^{-c}$ and $2^{|\Lambda_N|}e^{c}$, it follows that $|p_N(0) - \log 2| \le \beta(dJ + \Gamma) =: b$; as $|p_N'| \le \beta$, also $|p_N(x) - \log 2| \le b + \beta|x|$ for real x.

Fix $x_0 > 0$, and let $R = x_0/2$ and $r = x_0/4$. On the disc $|h - x_0| \le R$ we have $\operatorname{Re} f_N \le \log 2 + b + \frac{3}{2}\beta x_0$, and $f_N(x_0) = p_N(x_0) \ge \log 2 - b - \beta x_0$. By the Borel–Carathéodory inequality (see Fact P.55(c)),

$$|f_N(h) - f_N(x_0)| \le \frac{2r}{R-r}\Big(\sup_{|z-x_0|\le R} \operatorname{Re} f_N(z) - \operatorname{Re} f_N(x_0)\Big) \le 2\big(2b + \tfrac{5}{2}\beta x_0\big)$$

for $|h - x_0| \le r$. Since also $|f_N(x_0)| \le \log 2 + b + \beta x_0$, the functions f_N are uniformly bounded on each disc $\{|h - x_0| < x_0/4\}$, and therefore locally uniformly bounded on the connected open set U that is the union of these discs, which contains $(0,\infty)$. They converge on $(0,\infty)$, by (Q2). By Vitali's theorem (see Fact P.55(d)), they converge uniformly on compact subsets of U to a holomorphic function F, which equals p on $(0,\infty)$; hence the pressure p is real analytic there, and $m_N = f_N'/\beta \to F'/\beta = p'/\beta$ on $(0,\infty)$. □

We define the spontaneous magnetization of the quantum model by

$$m^*(\beta,\Gamma) = \frac{p'(0^+)}{\beta};$$

by evenness, $p'(0^-)/\beta = -m^*(\beta,\Gamma)$. With offsets as in the classical model, let $r_{N,a}(h) = m_N(h-a)$ for $N \ge 1$, $a \in [-1/\beta, 1/\beta]$ and $h \in Y = [-2/\beta, 2/\beta]$, and let $\mathcal{R}_\Gamma$ be the family of these functions. For $a \in [-1/\beta, 1/\beta]$ and $c \in [-m^*, m^*]$, where $m^* = m^*(\beta,\Gamma)$, let $g_{a,c}$ be defined as before Lemma P.53, with the pressure p of the quantum model.

Theorem P.58. *Let $d \ge 1$, $J > 0$, $\Gamma \ge 0$ and $\beta > 0$ be given, let $m^*(\beta,\Gamma)$ be the spontaneous magnetization of the quantum model, and let $\mathcal{R}_\Gamma$ be the family of responses $r_{N,a}$ of the model with offsets.*

(a) *If $m^*(\beta,\Gamma) = 0$, then every idealized response is continuous: the family $\mathcal{R}_\Gamma$ is stable, and the closure of the family is metrizable.*

(b) *If $m^*(\beta,\Gamma) > 0$, then the family $\mathcal{R}_\Gamma$ is unstable, and the closure of the family is first countable but not hereditarily separable: the family is NIP$^{\mathrm{fc}}$ but not NIP$^{\mathrm{hs}}$. For each*

$a \in (-1/\beta, 1/\beta)$, the functions g_{a,m^*} and $g_{a,-m^*}$ lie in the closure and differ only at a, and their midpoints $g_{a,0}$ form an uncountable discrete set.

The family $\{m_N\}$ without offsets is stable exactly when $m^*(\beta,\Gamma) = 0$, and is NIP$^{\mathrm{met}}$ in all cases.

Proof. The discussion preceding Lemma P.53, the lemma and the proof of Theorem P.54 use only the following properties: the function m_N is continuous, nondecreasing and odd, with $|m_N| \le 1$ and $m_N(0) = 0$; the limit p exists, is convex and even, and is differentiable at every $h \neq 0$, with $m_N(h) \to p'(h)/\beta$; and $p'(0^{\pm})/\beta = \pm m^*$. These hold by (Q1), (Q2) and Proposition P.57, and the proofs apply verbatim with $m^*(\beta,\Gamma)$ in place of $m^*(\beta)$. □

When $m^* > 0$, for each interior offset $a \in (-1/\beta, 1/\beta)$ the values $g_{a,\pm m^*}(a) = \pm m^*$ are the magnetizations of the two pure phases at the transition point $h = a$, and the value $g_{a,0}(a) = 0$ is that of the symmetric state. The classification distinguishes vanishing from nonvanishing spontaneous magnetization at positive temperature. It does not determine the critical parameters; those enter through the physical results recalled below. Both cases of the theorem occur, by results of Björnberg and Grimmett [**BG09**]: for $d = 1$ and $\beta < \infty$ the spontaneous magnetization vanishes, and for $d \ge 2$ and $\Gamma > 0$ there is a critical value of $2J/\Gamma$, depending on d and $\beta\Gamma$, below which $m^* = 0$ and above which $m^* > 0$ (see [**BG09**, Section 1 and Theorems 1.1–1.2], where the existence of the critical value is attributed to standard methods). Their conventions match ours as follows. They use the zero-field Hamiltonian $-\frac{1}{2}\lambda \sum \sigma^z\sigma^z - \delta \sum \sigma^x$, and, with a longitudinal term $-\gamma \sum \sigma^z$, the parameters correspond to $\lambda = 2J$, $\delta = \Gamma$ and $h = \gamma$. They define the spontaneous magnetization as the limit, as $\gamma \downarrow 0$, of the infinite-volume magnetization at field γ on the same periodic boxes Λ_N, through the space-time representation. By translation invariance and Proposition P.57, the infinite-volume magnetization is $p'(\gamma)/\beta$, and since the right derivative of a convex function is the right limit of the derivative, their spontaneous magnetization is $p'(0^+)/\beta = m^*(\beta,\Gamma)$. Order at low temperature and small Γ/J goes back to Ginibre [**Gin69**].

The Lee–Yang theorem [**LY52**] was extended to general classical ferromagnets by Newman [**New74**] and by Lieb and Sokal [**LS81**], and Lee–Yang theorems for quantum spin systems go back to Asano [**Asa70**] and to Suzuki and Fisher [**SF71**]. The formulation of Fröhlich and Rodriguez [**FR12**, Theorem 5], applied to H_N after relabeling the axes, with $h_1 = h$, $h_2 = \Gamma$ and $h_3 = 0$ in their condition $\operatorname{Re} h_1 > |h_2| + |h_3|$, gives the zero-free region $\operatorname{Re} h > \Gamma$. Proposition P.56 gives the whole half-plane $\operatorname{Re} h > 0$ for the transverse-field Ising model, by a Suzuki–Trotter approximation by classical ferromagnets followed by Hurwitz's theorem, and Proposition P.57 derives from the theorem the analytic input needed for the response classification; the argument uses standard quantum Lee–Yang methods, in the spirit of Simon and Griffiths [**SG73**] and of Suzuki [**Suz76**].

P.3. Part III: simplicity

In this subsection we prove the results of Part III that are not proved in the variations. The simplicity of the two theories in Theorem 22.1, with their nonforking independence, is proved in Theorems P.59 and P.61. In Theorem 29.1, part (a) is Theorem P.70, and part (b) is Theorem P.76 applied to the Rado observable, whose pure responses are freely independent by Lemma P.67; the affine conditions with bounded coefficients, in which the number of parameters may vary, are treated in Theorem P.74 and Corollary P.75.

The Rado graph and the Pauli group have simple unstable theories. We prove here that the theories of the Rado graph and of the Pauli group are simple and unstable, and that in both structures nonforking independence is the independence given by algebraic closure (see Theorems P.59 and P.61). The facts are known [**She80**, **CH03**], and we include the short proofs because the independence relations that they compute are the ones translated into the language of stabilizer codes in Variatio 23. Each structure is a Fraïssé limit whose theory eliminates quantifiers, so that types are determined by the atomic relations: edges and equalities in the Rado graph, linear relations and values of the symplectic form in the Pauli group. The difference between the two structures lies in algebraic closure, which is trivial in the Rado graph and is the linear span in the Pauli group.

The Rado graph. Recall from Variatio 21 that the Rado graph (R, E) is the Fraïssé limit of the finite graphs. A countable graph is isomorphic to the Rado graph if and only if it has the *extension property* in the following form: for all finite disjoint sets R_1 and R_0 of vertices, there are infinitely many vertices outside $R_1 \cup R_0$ adjacent to every vertex of R_1 and to no vertex of R_0 [**Fra54**, **Rad64**]. The theory of (R, E) is $\aleph_0$-categorical and eliminates quantifiers. Hence the type of a tuple over a set D is determined by the equalities and edges among the entries of the tuple and between those entries and the elements of D. By saturation, the monster model $\mathbb{M}$ of the theory of (R, E) has the extension property for small disjoint sets R_1 and R_0 of vertices. In particular, the algebraic closure of a small set D of vertices of $\mathbb{M}$ is D itself: for a vertex v outside D, the extension property, applied to the neighbors of v in D and to the remaining vertices of D, gives infinitely many vertices outside D with the same edges to D as v, and quantifier elimination gives them the type of v over D, so v is not algebraic over D.

The independence property of $E(x, y)$, and hence the instability of the theory, were shown in Variatio 21.

Theorem P.59. *The theory of the Rado graph is simple, and, for small sets A, B and C of vertices of the monster model $\mathbb{M}$ of this theory, nonforking independence $A \mathop{\smile\!\!\!\!\!\!|}_C B$ holds if and only if $A \cap B \subseteq C$.*

Proof. We verify the hypotheses of the Kim–Pillay theorem (see Variatio 20) for the relation $A \mathop{\smile\!\!\!\!\!\!|}_C B$ defined by $A \cap B \subseteq C$. The first six are set algebra, and extension and the independence theorem follow from the extension property of $\mathbb{M}$. Invariance, symmetry and finite character are immediate. Monotonicity and transitivity hold together, since the conditions $A \cap B \subseteq C$ and $A \cap D \subseteq C \cup B$ say jointly that every element of $A \cap (B \cup D)$ lies in C. For local character, given a finite tuple a and a set B, the finite set $B_0 = B \cap a$ satisfies $a \cap B \subseteq B_0$, which is the relation $a \mathop{\smile\!\!\!\!\!\!|}_{B_0} B$.

In the two constructions below, the coordinates whose entries lie in the base are listed first, with the same permutation of coordinates for two tuples that have the same type over the base, and the original order is restored at the end; repeated entries keep their pattern of equalities, since one new vertex is chosen for each distinct entry outside the base. For extension, let a and $C \subseteq B$ be given, and write $a = (c, a_0)$, where c lists the entries of a in C. The type of a_0 over C prescribes only the equalities and edges of a_0 within itself and to C. Hence the extension property of $\mathbb{M}$, applied to the distinct entries in turn, with $B \setminus C$ added to the set of vertices to which nonadjacency is prescribed, gives a tuple a_0' disjoint from B with the same type over C, and 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 the independence theorem over models (see Variatio 20). Since $\operatorname{tp}(a_1/M) = \operatorname{tp}(a_2/M)$, the two tuples have the same entries in M and the same remaining equalities; write $a_\ell = (m, a_\ell')$ for $\ell = 1, 2$, where the

tuple m lists these common entries and each a'_ℓ is disjoint from M. The hypotheses $a_1 \mathop{\smile\!\!\!\!\!\!|}_M B$ and $a_2 \mathop{\smile\!\!\!\!\!\!|}_M D$ say that a'_1 is disjoint from B and a'_2 from D. We look for a tuple x, disjoint from $M \cup B \cup D$, whose equalities and edges within itself and edges to M are those of a'_1 (equivalently, of a'_2), whose edges to $B \setminus M$ are those of a'_1, and whose edges to $D \setminus M$ are those of a'_2. These prescriptions concern disjoint sets of vertices, because $B \cap D \subseteq M$; the extension property of $\mathbb{M}$ therefore provides such an x, and by quantifier elimination the tuple $a = (m, x)$ is as required. □

Pauli strings. We use the labels $(x, z) \in \mathbb{F}_2^{2n}$ of Pauli strings, the commutation form ω and the space $V = \bigoplus_{k\in\mathbb{N}} \mathbb{F}_2^2$ of labels of the Pauli strings of finite support, the Pauli group (V, ω), of Variatio 22. For a label $v = (x, z)$ we write P_v for the Hermitian string $\mathrm{i}^{N(v)} X^{x_1} Z^{z_1} \otimes \cdots \otimes X^{x_n} Z^{z_n}$, where $N(v)$ is the number of indices k with $x_k = z_k = 1$, so that $P_v^2 = I$. The *syndrome* of an error e with respect to checks $g_1, \ldots, g_m$, for a positive integer m, is the vector $(\omega(g_1, e), \ldots, \omega(g_m, e))$.

The Pauli group as a Fraïssé limit. By Variatio 24, (V, ω) is the Fraïssé limit of the finite alternating spaces over $\mathbb{F}_2$, in the language with $+$, 0 and a binary relation for $\omega(x, y) = 1$. The theory of (V, ω) 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 $\mathrm{span}(D)$ and the values of ω on $\mathrm{span}(aD)$. The algebraic closure of a set is the span of the set, as we show after Lemma P.60.

Lemma P.60. *Let n be a positive integer. If $a_1, \ldots, a_n \in V$ are linearly independent and $\varepsilon_1, \ldots, \varepsilon_n \in \mathbb{F}_2$, then there is $b \in V$ with $\omega(a_i, b) = \varepsilon_i$ for $i = 1, \ldots, n$.*

Proof. The vectors a_i lie in the span W of the labels of X and Z on the first m qubits, for a large enough integer m, and the form is nondegenerate on W. Hence the functionals $\omega(a_i, \cdot)$ restricted to W are linearly independent, and the linear map $b \mapsto (\omega(a_i, b))_{i\le n}$ from W to $\mathbb{F}_2^n$ is onto. □

In the language of codes, Lemma P.60 says that every syndrome with respect to finitely many linearly independent checks is the syndrome of some error. Since V contains infinitely many linearly independent vectors, compactness then shows that the formula $\omega(x, y) = 1$ has the independence property, and the theory of (V, ω) is unstable.

We also use the finite-extension property of the Fraïssé limit. Every finite alternating extension of a finite subspace of V embeds into V over the subspace: embed the extension into V, then use homogeneity to match the given copy of the subspace. For each pair of finite alternating spaces, this property is expressed by a first-order extension sentence, so it holds in the monster model $\mathbb{M}$ as well. Consequently, for every small subspace D of $\mathbb{M}$ and every positive integer n, the conditions describing an alternating extension $D \oplus \mathrm{span}(x_1, \ldots, x_n)$ by n independent vectors are finitely satisfiable: each finite set of conditions concerns only a finite-dimensional subspace of D and a finite alternating extension of it. Saturation realizes these conditions over D. In particular, every linear functional f on D has the form $\omega(\cdot, b)$ on D for some $b \in \mathbb{M}$.

The same property computes algebraic closure. Let A be a small subset of $\mathbb{M}$, let $D = \mathrm{span}(A)$, and let v be a vector of $\mathbb{M}$ outside D. For a positive integer n, prescribe an alternating form on $D \oplus \mathrm{span}(x_1, \ldots, x_n)$ by requiring

$$\omega(d, x_i) = \omega(d, v) \quad (d \in D,\ 1 \le i \le n), \qquad \omega(x_i, x_j) = 0 \quad (1 \le i, j \le n).$$

The finite-extension property and saturation realize this space over D by vectors linearly independent over D. These vectors are distinct, and quantifier elimination gives each of them the type of v over D. Hence the type of v over A has infinitely many realizations, and

v is not algebraic over A; since every element of $\mathrm{span}(A)$ is the value of a term at elements of A, the algebraic closure of A is $\mathrm{span}(A)$.

Theorem P.61. *The theory of (V, ω) is simple, and, for small subsets A, B and C of the monster model $\mathbb{M}$ of this theory, nonforking independence $A \mathop{\smile\!\!\!\!\!\!|}_C B$ holds if and only if* $\mathrm{span}(AC) \cap \mathrm{span}(BC) = \mathrm{span}(C)$.

Proof. The proof follows that of Theorem P.59, with linear spans in place of the sets themselves: the first six axioms are linear algebra, and extension and the independence theorem come from the freedom to prescribe the form on a complement. We verify the hypotheses of the Kim–Pillay theorem (see Variatio 20) for the relation $A \mathop{\smile\!\!\!\!\!\!|}_C B$ defined by $\mathrm{span}(AC) \cap \mathrm{span}(BC) = \mathrm{span}(C)$. Invariance and symmetry are immediate, and finite character holds because a vector of $\mathrm{span}(AC) \cap \mathrm{span}(BC)$ outside $\mathrm{span}(C)$ is a combination of finitely many elements of A, B and C.

For monotonicity and transitivity, let $Q_1 = \mathrm{span}(AC)$, $Q_2 = \mathrm{span}(BC)$, $Q_3 = \mathrm{span}(BCD)$ and $Q_0 = \mathrm{span}(C)$; these subspaces satisfy $Q_2 \subseteq Q_3$ and $\mathrm{span}(ABC) = Q_1 + Q_2$. 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 $(Q_1 + Q_2) \cap Q_3 = Q_2$. Suppose that $Q_1 \cap Q_3 = Q_0$. Since $Q_2 \subseteq Q_3$, the intersection $Q_1 \cap Q_2$ is contained in Q_0. If $u_1 + u_2 \in Q_3$ with $u_1 \in Q_1$ and $u_2 \in Q_2$, then the vector u_1 also lies in Q_3, hence in $Q_1 \cap Q_3 = Q_0 \subseteq Q_2$; thus the sum lies in Q_2. Conversely, every $u \in Q_1 \cap Q_3$ lies in $(Q_1 + Q_2) \cap Q_3 = Q_2$, and therefore in $Q_1 \cap Q_2 = Q_0$.

Local character follows from finite dimension. Given a finite tuple a and a set B, the space $\mathrm{span}(a) \cap \mathrm{span}(B)$ is finite, so it is contained in $\mathrm{span}(B_0)$ for some finite $B_0 \subseteq B$. If $u + w \in \mathrm{span}(B)$ with $u \in \mathrm{span}(a)$ and $w \in \mathrm{span}(B_0)$, then the vector u lies in $\mathrm{span}(a) \cap \mathrm{span}(B) \subseteq \mathrm{span}(B_0)$. It follows that $\mathrm{span}(aB_0) \cap \mathrm{span}(B) = \mathrm{span}(B_0)$, which is the relation $a \mathop{\smile\!\!\!\!\!\!|}_{B_0} B$.

For extension and the independence theorem, we first treat finite tuples. For a small tuple, each finite subset of the required type and of the span-independence conditions involves only finitely many coordinates; the finite-tuple constructions therefore make these conditions finitely satisfiable, and compactness and saturation realize them together. In either construction, choose a maximal subtuple linearly independent over the base span. Every omitted coordinate is a linear combination of the chosen coordinates and base parameters, and is restored by the same equation after the construction. In the independence theorem, equality of the two types over the base ensures that the same coordinate indices can be chosen and that these restoring equations agree. For extension, let a and $C \subseteq B$ be given, with a now linearly independent over $\mathrm{span}(C)$. Choose a complement B' of $\mathrm{span}(C)$ in $\mathrm{span}(BC)$. Take a tuple x of new vectors of the same length as a and prescribe the alternating form on $\mathrm{span}(BC) \oplus \mathrm{span}(x)$. On $\mathrm{span}(xC)$ use the values prescribed by $\mathrm{tp}(a/C)$, and require x to be orthogonal to B'. The consequence of the finite-extension property established before the theorem realizes this space over $\mathrm{span}(BC)$ by a tuple a'. This tuple has the type of a over C and is linearly independent over $\mathrm{span}(BC)$. Hence $\mathrm{span}(a'C) \cap \mathrm{span}(BC) = \mathrm{span}(C)$, which is the relation $a' \mathop{\smile\!\!\!\!\!\!|}_C B$.

The independence theorem reduces to a gluing of functionals. Let M, B, D, a_1 and a_2 be as in the independence theorem over models (see Variatio 20). As in the proof of extension, we may assume that a_1 is linearly independent over M; then so is a_2, and the hypotheses $a_1 \mathop{\smile\!\!\!\!\!\!|}_M B$ and $a_2 \mathop{\smile\!\!\!\!\!\!|}_M D$ make a_1 linearly independent over $\mathrm{span}(MB)$ and a_2 over $\mathrm{span}(MD)$. Let n be the common length of a_1 and a_2, and, for $\ell = 1, 2$ and $k = 1, \ldots, n$, let $a_{\ell,k}$ be the kth entry of a_ℓ. For each k, the functionals $\omega(a_{1,k}, \cdot)$ on $\mathrm{span}(MB)$ and $\omega(a_{2,k}, \cdot)$ on $\mathrm{span}(MD)$ agree on M, because $\mathrm{tp}(a_1/M) = \mathrm{tp}(a_2/M)$. Since $B \mathop{\smile\!\!\!\!\!\!|}_M D$, their domains

meet in $\operatorname{span}(MB) \cap \operatorname{span}(MD) = M$, and the two functionals glue to a linear functional on $\operatorname{span}(MBD)$. Realize over $\operatorname{span}(MBD)$ a tuple a, linearly independent over $\operatorname{span}(MBD)$, with these functionals and with the values of ω on a given by $\operatorname{tp}(a_1/M)$. By quantifier elimination, the type of a over MB is that of a_1, and the type of a over MD is that of a_2; the relation $a \mathop{\smile\!\!\!\!\!\!|}_M BD$ holds by construction. □

Remark P.62. The proof of Theorem P.61 applies over every finite field $\mathbb{F}_q$, for a prime power q, in the language with a unary function for multiplication by each scalar and a binary relation for each value of the form. For symmetric forms in characteristic two, and for arbitrary bilinear forms, the identification of the limit through symplectic bases and Witt's theorem is no longer available, but the Fraïssé limit still exists, since the finite spaces with such a form have free amalgamation. The verification of the Kim–Pillay axioms goes through with the values $\omega(x, x)$ prescribed by the type and, for forms that are not symmetric, with the functionals $\omega(x, \cdot)$ and $\omega(\cdot, x)$ glued separately.

In both structures, then, nonforking independence is the independence given by algebraic closure: for small sets A, B and C, the relation $A \mathop{\smile\!\!\!\!\!\!|}_C B$ holds exactly when the closures of AC and BC meet in the closure of C. The independence theorem reduces to the freedom to prescribe the remaining relations, edges or values of the form, on each side separately. In the Pauli group, closures are spans of labels and the prescribed values of the form are syndrome bits, and in Variatio 23 these facts become a dictionary between the model theory of (V, ω) and stabilizer codes.

Observables, mixtures and dividing. We now pass from a relation on labels to an observable and its mixed preparations. This changes the structure whose theory is being studied: a probability distribution on labels becomes one preparation, and the response records its average against a probe. Although the Pauli and Rado label theories are both simple and unstable, their observables have different probe spaces: linear functionals on the Pauli space and arbitrary sets of Rado vertices describe the pure probes. We construct the two response structures and formulate the tree properties for their real-valued predicates, preparing the comparison of the two observables in sections P.3 and P.3.

Diagonal observables. Let A and B be countable sets, let $\varphi\colon A \times B \to \{0, 1\}$ be a function, and recall from Variatio 28 the diagonal observable C_φ on $\ell^2(A) \otimes \ell^2(B)$, which reads $\varphi(a, b)$ on the pair of labels (a, b). A density operator of Alice enters only through its diagonal, a probability measure μ on A, a *mixture* of her labels, and its effective observable (see Variatio 11) is the operator of multiplication by the *actual response*

$$r_\mu(b) = \sum_{a \in A} \mu(a)\varphi(a, b) \qquad (b \in B).$$

The probe algebra of Variatio 11 is therefore the commutative C*-algebra of bounded functions on B generated by the constants and the indicators $b \mapsto \varphi(a, b)$, $a \in A$, and the probe space is its state space, with the weak* topology. Both are described through the closure K_φ of the set of functions $\varphi(\cdot, b)$, $b \in B$, in $\{0, 1\}^A$, a compact metrizable space, on which Radon and Borel probability measures coincide.

Lemma P.63. *Let A and B be countable sets, let $\varphi\colon A \times B \to \{0, 1\}$ be a function, and let K_φ be the closure of the set of functions $\varphi(\cdot, b)$, $b \in B$, in $\{0, 1\}^A$. Then the probe algebra of the diagonal observable C_φ is isomorphic to the algebra $C(K_\varphi)$ of continuous complex functions on K_φ, the probe space is the space $\mathcal{P}(K_\varphi)$ of Radon probability measures on K_φ with the weak* topology, and the completed response of a mixture μ is the affine continuous function*

$\mathfrak{r}_\mu\colon \mathcal{P}(K_\varphi) \to [0,1]$ *given by*

$$\mathfrak{r}_\mu(\nu) = \int_{K_\varphi} \sum_{a\in A} \mu(a)\lambda(a)\, d\nu(\lambda) \qquad (\nu \in \mathcal{P}(K_\varphi)).$$

Proof. For each $a \in A$, let q_a be the coordinate function $q_a(\lambda) = \lambda(a)$ on K_φ. These functions are real-valued and separate points of K_φ. The unital complex algebra they generate is therefore closed under complex conjugation and uniformly dense in $C(K_\varphi)$ by the Stone–Weierstrass theorem. Restriction along $b \mapsto \varphi(\cdot, b)$ preserves the supremum norm, because this map has dense image in K_φ, and sends q_a to the indicator $b \mapsto \varphi(a, b)$. It thus identifies $C(K_\varphi)$ with the probe algebra.

By the Riesz representation theorem, a state of $C(K_\varphi)$ is integration against a unique Radon probability measure; pointwise convergence on $C(K_\varphi)$ is precisely the stated weak* topology on these measures. For a mixture μ, the series $\sum_a \mu(a)q_a$ converges uniformly, since the norm of a tail is bounded by the corresponding probability mass. This function corresponds to the effective observable of the preparation with diagonal μ. Integrating it against the measure ν gives the displayed response formula, which is affine and weak* continuous in ν. □

The pure probes are the point masses δ_λ, $\lambda \in K_\varphi$, and we identify each of them with its point λ; a pure probe is thus an *idealized label* of Bob, recording which of Alice's labels it answers with 1. The actual labels, the functions $\varphi(\cdot, b)$ for $b \in B$, are dense among them, since K_φ is the closure of the set of these functions. For $\lambda \in K_\varphi$, we write $\mathfrak{r}_\mu(\lambda)$ as shorthand for $\mathfrak{r}_\mu(\delta_\lambda)$, the completed response at the corresponding pure probe. For every actual label $b \in B$, the completed response recovers the actual response:

$$\mathfrak{r}_\mu(\delta_{\varphi(\cdot,b)}) = r_\mu(b).$$

The response r_μ is defined on B, whereas the completed response $\mathfrak{r}_\mu$ is defined on the probe space, which contains the point masses of the actual labels.

The structure of an observable. The theory of the observable concerns its mixtures and probes, rather than its label structure alone. For a diagonal observable C_φ, let $\mathcal{M}_\varphi$ be the following two-sorted structure. Its preparation sort consists of Alice's probability mixtures on A, and its probe sort is $\mathcal{P}(K_\varphi)$. Its only predicate is the response $R(\mu, \nu) = \mathfrak{r}_\mu(\nu)$ with values in $[0, 1]$. As in Variatio 11, we use Keisler's model theory for real-valued structures [**Kei23**], without prescribing a metric in advance. We call C_φ *simple* when the real-valued theory of $\mathcal{M}_\varphi$ is simple in the sense of Ben Yaacov [**BY03**]. An elementary extension is a larger structure in which every formula with parameters from $\mathcal{M}_\varphi$ retains its value; its new elements of the preparation sort are called idealized preparations. For the Pauli observable, the error labels of arbitrary support already occur as pure probes in the original structure, as we show in Proposition P.66, and elementary extension is a further enlargement. Every tree property of the theory of $\mathcal{M}_\varphi$ is witnessed in an $\aleph_1$-saturated elementary extension, where $\aleph_1$-saturation means that a countable set of conditions with parameters is realized whenever each finite subset can be satisfied within every positive error. We say that actual preparations witness the property when the preparation parameters of a witnessing array lie in $\mathcal{M}_\varphi$ itself and the finite parts of its paths are realized in the original probe sort, in the sense of Definition P.64. For the Pauli observable, this does not require a whole infinite path to be realized by an error of finite support.

Dividing for families of functions. The passage from truth values to expectations expresses inconsistency through uniformly separated real values. As in Variatio 11, an independent sequence of responses realizes every finite high/low pattern at two fixed thresholds. For the tree properties, individual conditions are satisfiable, but fixed-size subfamilies within rows or sibling sets are uniformly inconsistent, while finite portions of each path remain jointly satisfiable. The common positive gap supplies the inequalities required in the real-valued formulation, and these finite configurations persist in elementary extensions of the response structure $\mathcal{M}_\varphi$. We state the function-family formulation before constructing the arrays of mixtures.

Definition P.64. Let Y be a set and $\mathcal{R}$ a family of real functions on Y.

(1) For an integer $k \geq 2$, 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 $j_1 < \cdots < j_k$ no $y \in Y$ satisfies $r_{j_\ell}(y) > \alpha$ for $\ell = 1, \ldots, k$.

(2) The family $\mathcal{R}$ has the *tree property of the second kind* (TP_2) 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 $f\colon \mathbb{N} \to \mathbb{N}$ and every $n \in \mathbb{N}$ there is $y \in Y$ with $r_{if(i)}(y) \geq \beta$ for all $i < n$.

The tree property of Variatio 20, its variant of the first kind, and TP_2 all require infinite sequences of instances that are k-inconsistent, for some integer $k \geq 2$: the children of a node, or a row. For a family of functions, we define the tree property and the tree property of the first kind from Variatio 20 in the same way as TP_2 in Definition P.64, with fixed real thresholds $\alpha < \beta$: the consistency of a branch is expressed by simultaneous values $\geq \beta$ of every finite set of functions along the branch, and the inconsistency of a specified finite set by the absence of a point at which all its functions have values $> \alpha$. In each of the three properties, the children of a node, or the members of a row, then form a uniformly k-inconsistent sequence, with $k = 2$ for the tree property of the first kind. A family that, for no integer $k \geq 2$, contains a uniformly k-inconsistent sequence therefore has none of the three tree properties. Applied to the family of responses of an observable, condition (2) of Definition P.64 says that actual preparations witness the tree property of the second kind of the atomic formula, in the real-valued logic of Variatio 11; the gap between the constants α and β of Definition P.64 is what the definitions of dividing in real-valued logic require [**BY03**].

The Pauli observable. The *Pauli observable* is the diagonal observable of the symplectic form ω on $V \times V$, where V is the space of labels of the Pauli strings on countably many qubits that act nontrivially on only finitely many of them (see section P.3): Alice and Bob each hold a Pauli label, and the observable reports whether the two Pauli strings anticommute. We call Alice's label the check and Bob's label the error; the observable then reports 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 P.60, which provides an error with any prescribed syndrome with respect to finitely many linearly independent checks, the observable has the independence property already on pure preparations.

Lemma P.65. *For the Pauli observable, the space K_ω is the group V^* of all linear functionals $V \to \mathbb{F}_2$.*

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 P.60 gives e with $\omega(\cdot, e)$ in the neighborhood. □

The functionals of V^* have a direct physical meaning. Let $\hat{V} = \prod_{k\in\mathbb{N}} \mathbb{F}_2^2$, with the product topology, be the group of labels of Pauli strings of arbitrary support, the support of a string being the set of qubits on which it acts nontrivially; 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 P.3, which is a finite sum because a has finite support. For p with $0 < p \le 1$, 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 Radon probability measure on $\hat{V}$ specifies a *Pauli channel* on the norm closure of the algebra of operators supported on finitely many qubits: on each finite set of qubits, the channel averages conjugation by the Pauli strings given by the restrictions of the random label to that set. We identify the channel with this probability law. An arbitrary-support label thus specifies compatible actions on all finite sets of qubits; no infinite tensor product unitary on a chosen Hilbert space is required.

Proposition P.66. *Let $\hat{V}$ be the group of labels of Pauli strings of arbitrary support, and for $0 < p \le 1$ 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*

$$\mathfrak{r}_{\delta_a}(\nu_p) = \tfrac{1}{2}\big(1 - (1 - \tfrac{4p}{3})^w\big).$$

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 $u = (x,z)$ with $z_k = \lambda(X_k)$ and $x_k = \lambda(Z_k)$ for every 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 P.63, Lemma P.65 and (a), the probes are the Radon probability measures on $\hat{V}$, and $\mathfrak{r}_\mu(\nu) = \int \sum_a \mu(a)\,\omega(a,u)\,d\nu(u)$, 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 second Borel–Cantelli lemma, almost every error acts nontrivially on infinitely many qubits. At each of the w qubits where the string with label a acts nontrivially, its factor anticommutes with two of the three nontrivial Pauli matrices, hence with the factor of 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 $\frac{1}{2}\big(1 - (1 - \frac{4p}{3})^w\big)$. □

In what follows we identify V^* with $\hat{V}$ through Proposition P.66(a) and call its elements *idealized errors*. By Proposition P.66(c), the depolarizing noise ν_p of each rate p, which models the noise of a memory on countably many qubits, is a probe of the Pauli observable concentrated on idealized errors that are not actual: with probability one, an error drawn from ν_p acts on infinitely many qubits. Alice's checks see such an error through finitely many

qubits at a time, and the response of a check of weight w, one that acts nontrivially on w qubits, tends to $\frac{1}{2}$, the value of a fair coin, as w grows. For $a \in V$, let χ_a be the character of V^* given by $\chi_a(\lambda) = (-1)^{\lambda(a)}$, a continuous homomorphism of V^* into the multiplicative group $\{\pm 1\}$. Alice's pure responses are the characters of this compact group up to an affine change: the pure response $\mathfrak{r}_{\delta_a}$ of a label a takes the value $\lambda(a) = (1-\chi_a(\lambda))/2$ at an idealized error λ.

The Rado observable. The *Rado observable* is the diagonal observable of the edge relation E on $R \times R$, where (R, E) is the Rado graph of section P.3: 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 of the Rado graph, which provides a vertex adjacent to every vertex of a finite set and to none of another finite set disjoint from it, the observable has the independence property already on pure preparations. We say that the diagonal observable C_φ of a function $\varphi\colon A \times B \to \{0,1\}$ 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 P.67. *For the Rado observable, the space K_E is the whole Cantor cube $\{0,1\}^R$. Hence the Rado observable has freely independent pure responses, while the Pauli observable does not.*

Proof. A basic open subset of $\{0,1\}^R$ prescribes adjacency to the vertices of a finite set R_1 and nonadjacency to those of a finite set R_0, and the extension property gives a vertex outside $R_1 \cup R_0$ with that pattern. By Lemma P.65, the space K_ω is the group V^*, which is a proper subset of $\{0,1\}^V$. □

The pure probes of the Rado observable are arbitrary sets of neighbors, that is, arbitrary subsets of R, 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 a set of neighbors drawn, independently, with law ν; it depends on ν only through the marginals $\nu\{\lambda : \lambda(a) = 1\}$, for $a \in R$, the probabilities that the individual sites 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. In sections P.3 and P.3 we show what this difference does to the tree property.

The pure and the mixed responses are NIP together. The tree property of the second kind in Theorem 28.1 is built from an independent sequence of actual labels, whereas NIP for the observable concerns all mixed responses on the probe space; the lemma below identifies the two conditions. We recall the fixed-gap definition of Variatio 11 for a family $\mathcal{R}$ of real functions on a set Y. A sequence $(r_n)_{n\in\mathbb{N}}$ in $\mathcal{R}$ is independent if there are real thresholds $\alpha < \beta$ with the following property. For all disjoint finite sets $N_0, N_1 \subseteq \mathbb{N}$, some $y \in Y$ satisfies $r_n(y) \le \alpha$ for $n \in N_0$ and $r_n(y) \ge \beta$ for $n \in N_1$. The family has the independence property when it contains such a sequence, and is NIP otherwise. For a diagonal observable C_φ of a function $\varphi\colon A \times B \to \{0,1\}$, recall from Lemma P.63 that K_φ is the closure of $\{\varphi(\cdot, b) : b \in B\}$ in $\{0,1\}^A$ and that the probe space is $\mathcal{P}(K_\varphi)$. For $a \in A$, let $f_a(\lambda) = \lambda(a)$ for $\lambda \in K_\varphi$, and, for a probability distribution μ on A, let $g_\mu = \sum_{a\in A} \mu(a) f_a$; by Lemma P.63, the response of μ is $\mathfrak{r}_\mu(\nu) = \int g_\mu \, d\nu$ for $\nu \in \mathcal{P}(K_\varphi)$.

Lemma P.68 (NIP for pure and mixed responses). *Let A and B be countable sets and let C_φ be the diagonal observable of a function $\varphi\colon A \times B \to \{0,1\}$. Then the following are equivalent:*

(i) *the family $\{\varphi(a,\cdot) : a \in A\}$ is NIP on B;*
(ii) *the family $\{f_a : a \in A\}$ of pure responses is NIP on K_φ;*

(iii) *the family* $\{\mathfrak{r}_\mu : \mu \text{ is a probability distribution on } A\}$ *of mixed responses is NIP on the probe space* $\mathcal{P}(K_\varphi)$.

PROOF. The functions f_a take only the values 0 and 1, so independence for them is equivalent to realizing every finite pattern of these two values. Each such pattern specifies a clopen subset of K_φ. If it is nonempty, it meets the dense set of actual labels $\{\varphi(\cdot, b) : b \in B\}$. Conversely, every pattern realized by an actual label is realized in K_φ. This proves (i)⇔(ii). If the pure responses are independent, the mixed responses associated with the point masses on the same labels are independent, by evaluation at the pure probes δ_λ. Thus (iii) implies (ii).

Assume (ii). Rosenthal's combinatorial dichotomy for uniformly bounded sequences of real functions [**Ros74**] gives every sequence of pure responses a subsequence that converges pointwise on K_φ, since an independent subsequence is excluded. This subsequence is weakly Cauchy in the real Banach space $C(K_\varphi)$ of continuous real-valued functions with the supremum norm: its values under each continuous linear functional form a Cauchy sequence. Indeed, every such functional is integration against a finite signed Radon measure, and the functions are bounded by 1, so dominated convergence applies. Hence $\{f_a : a \in A\}$ is weakly precompact, meaning that it is bounded and every sequence in it has a weakly Cauchy subsequence.

As in section P.2, we use Stegall's theorem: the absolutely convex hull of a weakly precompact set is weakly precompact (see [**Rod22**]). Here the absolutely convex hull consists of the finite sums $\sum_j c_j f_{a_j}$ with $\sum_j |c_j| \le 1$. Its norm closure is also weakly precompact: given a sequence $(h_n)_{n\in\mathbb{N}}$ in the closure, choose functions q_n in the hull with $\|h_n - q_n\|_\infty < 2^{-n}$ for $n \in \mathbb{N}$ and a weakly Cauchy subsequence of (q_n); the corresponding subsequence of (h_n) is weakly Cauchy as well. Each g_μ belongs to this closure, because its finite partial sums lie in the hull and converge uniformly, with the omitted probability mass bounding the error.

Every sequence (g_{μ_n}) therefore has a weakly Cauchy subsequence. Since integration against each probe $\nu \in \mathcal{P}(K_\varphi)$ is a continuous linear functional on $C(K_\varphi)$, the corresponding subsequence of responses $(\mathfrak{r}_{\mu_n})$ converges pointwise on the probe space. An independent sequence of these continuous responses cannot have a pointwise convergent subsequence. To see this, prescribe alternating low and high values along any subsequence at its fixed thresholds. Every finite prescription is realized, and the associated closed subsets of the compact probe space have the finite intersection property; hence one probe realizes the entire alternating prescription, preventing convergence there. Thus the mixed family is NIP, proving (iii). □

The Pauli observable: syndrome sectors and the first check to fire. For the Pauli observable, the linear relations among check labels allow an infinite array of actual mixtures to witness the tree property of the second kind for the response predicate itself. The construction uses the indicator of a syndrome sector, the set of idealized errors with specified commutation values against finitely many checks. Within each block of linearly independent checks, Alice draws a label uniformly from the coset of a chosen check modulo the span of its predecessors. The response is 1 when that check is the first to anticommute with Bob's error, and 1/2 when an earlier check already anticommutes with it. The resulting row conditions are incompatible at a fixed gap even for mixed probes, whereas choices from finitely many rows are realized by a finite-support error when all the checks together are linearly independent. Thus the array parameters are actual preparations; an entire infinite path is realized by a pure probe already in the original probe space, whose error label may have infinite support, as specified in Theorem P.70(a).

Syndrome sectors. Let d be a positive integer, let $g_1, \dots, g_d$ be linearly independent checks in the space V of labels of Pauli strings of finite support, with 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 $\Sigma_\sigma = \{\lambda \in V^* : \lambda(g_i) = \sigma_i \text{ for } i \le d\}$ of idealized errors with syndrome σ its *syndrome sector*. For $s \in G$ we write $\mathfrak{r}_s$ for the pure response $\mathfrak{r}_{\delta_s}$, the response of the point mass at s, and we recall that χ_s is the character of V^* given by $\chi_s(\lambda) = (-1)^{\lambda(s)}$.

Proposition P.69. *Let d be a positive integer, let $g_1, \dots, g_d \in V$ be linearly independent checks with span G, and let $\sigma \in \mathbb{F}_2^d$ be a syndrome, regarded as the linear functional on G with $\sigma(g_i) = \sigma_i$ for $i \le d$. Then the indicator of the syndrome sector Σ_σ satisfies, at every pure probe $\lambda \in V^*$,*

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

Consequently, if $\sigma \neq 0$, then $\mathbf{1}_{\Sigma_\sigma} = \mathfrak{r}_{\mu_\sigma^1} - \mathfrak{r}_{\mu_\sigma^0}$ 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 $\mathbf{1}_{\Sigma_0} = 1 - 2\mathfrak{r}_{\mu_G}$, 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 $\mathfrak{r}_{\mu_\sigma^1}(\nu) - \mathfrak{r}_{\mu_\sigma^0}(\nu)$ if $\sigma \neq 0$, and $1 - 2\mathfrak{r}_{\mu_G}(\nu)$ 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 $\chi_s = 1 - 2\mathfrak{r}_s$. 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 $\mathfrak{r}_{\mu_\sigma^1} - \mathfrak{r}_{\mu_\sigma^0}$. If $\sigma = 0$, the first sum is 2^d, and the second is $2^d \mathfrak{r}_{\mu_G}$. Both sides of each identity are affine in the probe, and integration against ν, together with Proposition P.66(b), yields the last statement. □

When the checks commute pairwise and carry signs, the operator $2^{-d} \sum_{s \in G} (-1)^{\sigma(s)} g_s$, where g_s is the element with label s of the group generated by the signed checks, is the projector onto the states of syndrome σ [Got97], and the probability of the outcome σ when the checks are measured on a code state corrupted by a Pauli channel ν is given by Proposition P.69. 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.

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

Theorem P.70 (The first check to fire). *Let $(g_{im})_{i,m \in \mathbb{N}}$ be linearly independent vectors of V, let $W_{ij} = \operatorname{span}(g_{im} : m < j)$ for $i, j \in \mathbb{N}$, 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 $f \colon \mathbb{N} \to \mathbb{N}$ and every $n \in \mathbb{N}$, 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 $\mathfrak{r}_{\mu_{if(i)}}(\lambda) = 1$ for all i;*

(b) *for every i, every $j \neq j'$ and every probe $\nu \in \mathcal{P}(V^*)$, we have $\min\big(\mathfrak{r}_{\mu_{ij}}(\nu), \mathfrak{r}_{\mu_{ij'}}(\nu)\big) \le 2/3$, and the bound $2/3$ is attained.*

Consequently, the family of responses of Alice has the tree property of the second kind in the sense of Definition P.64, 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

$$\mathfrak{r}_{\mu_{ij}}(\lambda) = \begin{cases} 1 & \text{if } \lambda(g_{im}) = 0 \text{ for } m < j \text{ and } \lambda(g_{ij}) = 1, \\ 1/2 & \text{if } \lambda(g_{im}) = 1 \text{ for some } m < j, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

In words, the response is 1 exactly when g_{ij} is the first check of the block $(g_{im})_{m\in\mathbb{N}}$, which we call block i, to fire, that is, to take the value 1 under λ.

(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 P.60 gives an error e with $\omega(\cdot, e)$ satisfying them, and by (3) the response $r_{\mu_{if(i)}}(e)$ equals 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 (3), the pair $\big(\mathfrak{r}_{\mu_{ij}}(\lambda), \mathfrak{r}_{\mu_{ij'}}(\lambda)\big)$ 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. Every listed pair (s, t) satisfies $s + 2t \le 2$, and the same inequality holds throughout their convex hull. Hence $3\min(s, t) \le s + 2t \le 2$, so the minimum of the two coordinates is at most $2/3$. Equality requires $s = t = 2/3$, and this point is $\frac{2}{3}(1, \frac{1}{2}) + \frac{1}{3}(0, 1) = (\frac{2}{3}, \frac{2}{3})$. This point is the pair of responses at the probe $\frac{2}{3}\delta_{\lambda_1} + \frac{1}{3}\delta_{\lambda_2}$, where λ_1 and λ_2 are idealized errors under which g_{ij} and $g_{ij'}$, respectively, are the first checks of block i to fire; such idealized errors exist by (a).

Statements (a) and (b) are the two clauses of Definition P.64 for the array $(\mathfrak{r}_{\mu_{ij}})$, with $k = 2$, $\alpha = 2/3$ and $\beta = 1$. □

Remark P.71. For qudits of prime dimension p, the quantum systems with Hilbert space $\mathbb{C}^p$, 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, the pth root of unity by which the products of the two strings in the two orders differ, is nontrivial and 0 otherwise. The proof of Theorem P.70 adapts as follows. A nonzero functional on a finite $\mathbb{F}_p$-space takes each value equally often, so in (3) the conditions $\lambda(g_{ij}) = 1$ and $\lambda(g_{im}) = 1$ become $\lambda(g_{ij}) \neq 0$ and $\lambda(g_{im}) \neq 0$, and the response $\frac{1}{2}$ becomes $\vartheta = \frac{p-1}{p}$. For $j < j'$, the pair of responses at a pure probe takes one of the values $(1, \vartheta)$, (ϑ, ϑ), $(0, \vartheta)$, $(0, 1)$ and $(0, 0)$. The point (ϑ, ϑ) lies below the segment from $(1, \vartheta)$ to $(0, 1)$, since $\vartheta < 1 - \vartheta + \vartheta^2$, 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 P.70(b) becomes $\frac{p}{p+1}$, and it is attained.

In the notation of Theorem P.70 and the language of Proposition P.69, 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 $\mathfrak{r}_{\mu_{ij}} - \mathfrak{r}_{\mu^0_{ij}}$ on pure probes, where μ^0_{ij} is the uniform probability measure on W_{ij}. Only the

linear independence of the checks enters Theorem P.70; 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 P.72. Index the qubits by pairs $(i,m) \in \mathbb{N}^2$ and call the qubits (i,m) for $m \in \mathbb{N}$ the chain i. Let $X_{i,m}$ and $Z_{i,m}$ denote the bit flip and phase flip on the qubit (i,m). The check $g_{im} = Z_{i,m}Z_{i,m+1}$ compares the neighboring qubits m and $m+1$ in the bit-flip repetition code on chain i. These checks commute pairwise and are linearly independent. Let $e \in \hat{V}$ be a bit-flip label with exponents $x_{i,m} \in \mathbb{F}_2$ for $(i,m) \in \mathbb{N}^2$. We write $e = \prod_{i,m} X_{i,m}^{x_{i,m}}$ as notation for this label, allowing infinitely many nonzero exponents; no infinite product of operators on a fixed Hilbert space is asserted. For the finite-support check g_{im} we have $\omega(g_{im}, e) = x_{i,m} + x_{i,m+1}$. This equals 1 exactly when one of the neighboring qubits is flipped and the other is not, giving a domain wall between them. By (3), the response $\mathfrak{r}_{\mu_{ij}}(e)$ of the uniform probability measure μ_{ij} on the coset of g_{ij} modulo the span of $g_{i0}, \dots, g_{i,j-1}$, as in Theorem P.70, is 1 exactly when the first domain wall of e on the chain i lies between the qubits j and $j+1$. For this code, Theorem P.70 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 of every qubit of the 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 P.73. The mechanism behind Theorem P.70 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, because $2\mathfrak{r}_a\mathfrak{r}_b = \mathfrak{r}_a + \mathfrak{r}_b - \mathfrak{r}_{a+b}$ for the pure responses $\mathfrak{r}_a = \mathfrak{r}_{\delta_a}$, $a \in V$, restricted to pure probes. By Proposition P.69, 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. In the first-order structure (V, ω), by contrast, a formula speaks about a bounded number of labels, and, by Theorem P.61, no formula has the tree property.

The coset mixtures of Theorem P.70 thus strengthen the construction of Theorem 28.1: the response predicate itself witnesses the tree property of the second kind, with actual preparations as parameters, because a preparation of $\mathcal{M}_\omega$ can average over a finite set of labels whose size grows along the array, whereas a formula of the label structure (V, ω) speaks about a bounded number of labels.

Table 8 collects, for the Pauli observable, the translations of sections P.3, P.3 and P.3. It concerns the real-valued structure $\mathcal{M}_\omega$ of preparations and probes of section P.3, for the predicate ω, which has probability distributions on V as preparations, and Radon probability measures on $\hat{V} = \prod_{k\in\mathbb{N}} \mathbb{F}_2^2$ as probes (see Proposition P.66). The compact space $\hat{V}$ consists of arbitrary-support error labels and contains V; its point masses are the pure probes, with labels in V giving actual finite-support errors. The symplectic pairing of a finite-support check with any label in $\hat{V}$ is a finite sum, so every response is defined (see Proposition P.66). A specified nonzero syndrome is a linear functional on the finite check span; its two level sets, the kernel and the complementary coset, are labeled by the values 0 and 1.

The Rado observable: marginals and idealized preparations. For the Rado observable, the theory of mixed preparations and probes has the tree property of the second kind, as for the Pauli observable, but the preparation parameters required to witness it are different. The pure probe space is the full cube of adjacency patterns, and the response of an actual mixture depends only on the probabilities with which a probe declares individual vertices to

Response structure	Pauli interpretation	Result
actual error labels $e \in V$	point masses at finite-support errors	Proposition P.66
all pure probes	point masses at arbitrary-support labels in $\hat{V}$	Proposition P.66
all probes	Pauli channels: probability laws on $\hat{V}$	Proposition P.66
response of a mixture	anticommutation probability for independent draws of a check and an error	Proposition P.66
uniform response on the 1-coset minus that on the 0-coset	probability of the specified nonzero syndrome	Proposition P.69
commuting signed checks	operator analogue: projector onto a syndrome eigenspace	Proposition P.69
TP_2 parameters	actual finite mixtures of checks	Theorem P.70
path realizations	finite path: an error in V; whole path: a pure probe in $\hat{V}$	Theorem P.70

TABLE 8. From response functions to syndrome probabilities: the response structure and its Pauli interpretation.

be neighbors. The freedom to prescribe these marginals separately on disjoint blocks prevents an infinite uniformly inconsistent row with actual mixture parameters, first for bounded affine combinations and then for each fixed continuous connective applied to finitely many responses (see Theorems P.74 and P.76). Consequently, a witnessing array for a quantifier-free formula of this form, in a probe variable with preparation parameters, must contain preparation parameters from an elementary extension outside the original structure. The restriction concerns the parameters of the array; the probes in the compatibility results range over the whole probe space, including mixed probes and pure probes that are not actual labels. The same argument applies to the random bipartite graph, whose two vertex classes serve as Alice's and Bob's labels.

Affine combinations of responses. Consider a diagonal observable with freely independent pure responses, that is, one for which every pattern $\lambda \in \{0,1\}^A$ of answers to Alice's labels, the elements of A, is an idealized label of Bob. The response of a mixture depends on a probe ν only through its marginals $\pi_a(\nu) = \nu\{\lambda : \lambda(a) = 1\}$, for $a \in A$, as for the Rado observable in section P.3, 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

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

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 conflict only through the labels on which both place weight. A gliding hump, in the proof of Theorem P.74 below, shows that a sequence of such functions along which $|\gamma| + \|c\|_1$ stays bounded has a subsequence whose members place their weight, up to small errors, on a common finite set of labels and on pairwise disjoint finite blocks; for every positive integer k, a single pure probe then nearly maximizes k of them.

Theorem P.74 (Compatibility of affine conditions). *Let C_φ be the diagonal observable of a function $\varphi\colon A \times B \to \{0,1\}$ on countable sets A and B, with freely independent pure responses, and let $\kappa > 0$ be given. Let $\mathcal{R}_\kappa$ be the family of functions $\gamma + \sum_{i<n} c_i \mathfrak{r}_{\mu_i}$ on the*

probe space $\mathcal{P}(\{0,1\}^A)$ of Radon probability measures on $\{0,1\}^A$, where $n \in \mathbb{N}$, the μ_i are mixtures of Alice, and γ and the c_i are reals with $|\gamma| + \sum_i |c_i| \leq \kappa$. Then the family $\mathcal{R}_\kappa$ contains no uniformly k-inconsistent sequence, for any integer $k \geq 2$. In particular, no tree property is witnessed in $\mathcal{R}_\kappa$.

Proof. Let $(h_j)_{j\in\mathbb{N}}$ 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 integer $k \geq 2$ there are $j_1 < \cdots < j_k$ and a pure probe λ with $h_{j_\ell}(\delta_\lambda) > \alpha$ for $\ell = 1, \ldots, k$.

Write $h_j = \gamma_j + \sum_i c_{j,i}\mathfrak{r}_{\mu_{j,i}}$, with real coefficients $c_{j,i}$ and mixtures $\mu_{j,i}$ of Alice; then $h_j = h_{c_j,\gamma_j}$ with $c_j = \sum_i c_{j,i}\mu_{j,i}$, whence $\|c_j\|_1 \leq \kappa$. 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 $\|c\|_1 \leq \kappa$. Let $\eta = (\beta - \alpha)/5$, and choose a finite set $F \subseteq A$ with $\|c \restriction (A \setminus F)\|_1 < \eta$. Let $F_0 = F$, and choose inductively indices $j_1 < \cdots < j_k$ and finite sets $F_0 \subseteq F_1 \subseteq \cdots \subseteq F_k$ such that

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

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 $\|c_{j_\ell} \restriction (F_{\ell-1} \setminus F)\|_1 \leq \|(c_{j_\ell} - c) \restriction F_{\ell-1}\|_1 + \|c \restriction (A \setminus F)\|_1 < 2\eta$. On $A \setminus F_\ell$ it is less than η. Hence $h_{j_\ell}(\delta_\lambda) > \sup h_{j_\ell} - 4\eta \geq \beta - 4\eta > \alpha$. □

Corollary P.75. *Let $n \in \mathbb{N}$ be given, and let γ and c_i, for $i < n$, be reals. For the Rado observable, no tree property of the affine formula $\gamma + \sum_{i<n} c_i R(y_i, x)$, in which R is the response predicate, the variables y_i are of the sort of preparations and the variable x is of the sort of probes, is witnessed by actual preparations, although the real-valued theory of the structure $\mathcal{M}_E$ of its mixed preparations and probes has the tree property of the second kind.*

Proof. By Lemma P.67, the Rado observable has freely independent pure responses. With actual preparations as parameters, the instances of the formula lie in the family $\mathcal{R}_\kappa$ of Theorem P.74 for every $\kappa > 0$ with $\kappa \geq |\gamma| + \sum_{i<n} |c_i|$, and the theorem yields the first statement. The second statement is Theorem 28.1. □

The distinction between finite and infinite arrays is essential here. For the fixed difference-of-responses formula of Theorem 28.1, actual mixtures provide arrays of every finite size, but Theorem P.74 excludes an infinite uniformly inconsistent row whose preparation parameters are all actual mixtures. In fact, in every row of a witnessing array for this formula, all but finitely many entries involve an idealized preparation: otherwise an infinite subsequence with both preparation parameters actual would contradict Theorem P.74. The arbitrarily large finite arrays therefore do not assemble into such a row inside the original structure. This conclusion also holds for the random bipartite graph, since its finite extension property makes every adjacency pattern a pure probe.

Quantifier-free formulas. In Theorem P.74 the conditions on a probe are given by affine combinations of responses. In the real-valued logic of Variatio 11, a quantifier-free formula in a variable x of the sort of probes, built from the response with parameters of the sort of preparations, 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 P.74 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 P.76 (Compatibility of quantifier-free conditions). *Let C_φ be the diagonal observable of a function $\varphi\colon A\times B\to\{0,1\}$ on countable sets A and B, with freely independent pure responses, let $n\geq 1$ be an integer, and let $u\colon[0,1]^n\to\mathbb{R}$ be a continuous function. Let $\mathcal{R}_u$ be the family of functions*

$$h(\nu)=u\big(\mathfrak{r}_{\mu_1}(\nu),\dots,\mathfrak{r}_{\mu_n}(\nu)\big)\qquad(\nu\in\mathcal{P}(\{0,1\}^A)),$$

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 integer $k\geq 2$. In particular, no tree property of the quantifier-free formula $u(R(y_1,x),\dots,R(y_n,x))$, where R is the response predicate, the variable x is of the sort of probes and the variables $y_1,\dots,y_n$ are of the sort of preparations, 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 $(\pi_a(\nu))_{a\in A}$, defined before Theorem P.74. Every $h\in\mathcal{R}_u$ factors as $h(\nu)=H(\pi(\nu))$ with

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

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 $(h_j)_{j\in\mathbb{N}}$ be a sequence in $\mathcal{R}_u$, with parameters $\mu_{j,1},\dots,\mu_{j,n}$ and associated functions H_j on the cube, and let $\alpha<\beta$ be such that each h_j takes a value $\geq\beta$. We show that for every integer $k\geq 2$ there are $j_1<\dots<j_k$ and $q\in[0,1]^A$ with $H_{j_\ell}(q)>\alpha$ for $\ell=1,\dots,k$. 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 $H_j(q^{(j)})\geq\beta$. The set A is countable and the cube is compact and metrizable, so a diagonal argument provides a subsequence, still indexed by j, along which $\mu_{j,i}\to\mu_i$ pointwise on A for each $i\leq n$ and $q^{(j)}\to q^\infty$ in $[0,1]^A$. By Fatou's lemma, each limit μ_i has total mass at most 1. Let $\eta=\theta/5$, and choose a finite set $F\subseteq A$ with $\mu_i(A\setminus F)<\eta$ for all $i\leq n$.

Let $F_0=F$, 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 $i\leq n$,

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

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 P.74; 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 $\mu_{j_\ell,i}(a) \le 1$. On $F_{\ell-1} \setminus F$ it is at most $\mu_{j_\ell,i}(F_{\ell-1} \setminus F) \le \sum_{a \in F_{\ell-1}} |\mu_{j_\ell,i}(a) - \mu_i(a)| + \mu_i(A \setminus F) < 2\eta$. 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 $i \le n$, and the choice of θ gives $H_{j_\ell}(q) > H_{j_\ell}(q^{(j_\ell)}) - (\beta - \alpha) \ge \alpha$. □

For affine combinations, the proof of Theorem P.74 yields slightly more: there the common witness can be taken pure. For an arbitrary connective u the witness is the product probe whose vector of marginals is the point q constructed in the proof of Theorem P.76, and it cannot always be taken pure: for $n = 1$, $\mu = \delta_a$ with $a \in A$, and $u(t) = 1 - |2t - 1|$, the instance vanishes at every pure probe and equals 1 at the probes ν with $\pi_a(\nu) = 1/2$.

Thus, for the Rado observable, the freedom to prescribe marginals on disjoint blocks of labels lets the gliding hump assemble compatible conditions from every sequence with actual parameters. For the Pauli observable, by contrast, the linear relations among label coordinates place the indicators of syndrome sectors in the affine span of the responses, with a coefficient bound independent of the number of checks (see Proposition P.69), and these identities yield the actual-parameter array of Theorem P.70.

Questions

Question 1. In classical model theory, a simple theory without the independence property is stable. Is a simple diagonal observable stable? By Theorem 28.1, a simple diagonal observable has no independence property, so the question is whether the unstable NIP diagonal observables, such as the energy comparison of the Aria, acquire a tree property once preparations are mixed.

Question 2. Which diagonal observables with the independence property have the tree property of the second kind witnessed by actual preparations? A criterion is suggested by Theorem 29.1: whether the linear span of the responses of Alice's labels, as functions on the space K_φ of Bob's idealized labels (see Variatio 28), contains, with coefficients of bounded total absolute value, the indicators of the sets of an infinite array of clopen sets whose rows consist of pairwise disjoint sets and in which every choice of one set from each of finitely many rows has nonempty intersection.

Question 3. The causal order of a $(1+1)$-dimensional spacetime is NIP but not $\mathrm{NIP}^{\mathrm{fc}}$: its idealized responses form a compactum that is not first countable (see Theorem 16.1). In which class of Todorčević's trichotomy does the causal order of a spacetime of higher dimension lie?

Question 4. The comparison of positions of a particle in a box realizes the minimal families D_3, D_4 and D_7 of the Argyros–Dodos–Kanellopoulos heptachotomy by trees whose supports shrink to the points of a Cantor set, and in these trees the departures on the two sides of a branch end in distinct limits (see Table 6). The family D_6, realized in Part II by a spin chain read both site by site and against templates, requires departures on the two sides that end in one common limit. Is there a tree of preparations of the comparison of positions, for instance one whose supports do not shrink to points, whose departures on the two sides of each branch share one limit and whose responses are equivalent to D_6?

Question 5. For qudits of prime dimension p, quantum systems with Hilbert space $\mathbb{C}^p$, the generalized Pauli strings, built from the shift $|s\rangle \mapsto |s+1\rangle$ and the clock $|s\rangle \mapsto e^{2\pi i s/p}|s\rangle$, for s modulo p, in place of X and Z, are labeled by a symplectic space over the field with p elements, and the dictionary of Variatio 23 carries over. For continuous variables, the commutation phase of a translation by x and a momentum kick by k is e^{ikx}, whose real part determines the interference fringe of Part II; for the wavenumbers 8^n the fringe is not NIP (see Theorem 16.1). Where do fringes of unbounded wavenumber lie with respect to simplicity?

Historical remarks

Shelah introduced stable theories in 1969 [She69], and the order property and the independence property of formulas in 1971 [She71]; simple theories followed in 1980, with the theory T_{ind} of the random graph as a motivating example [She80]. The double limit condition also has an analytic history: it is closely related to the regularity of bilinear forms studied by Arens in 1951 [Are51], it is Grothendieck's criterion of 1952 for weak compactness [Gro52], and it defines the stable Banach spaces introduced by Krivine and Maurey in 1981 [KM81]; Iovino connected it with Shelah's stability and developed the real-valued theory, including the stability spectrum for real-valued structures [Iov99a, Iov99b]. Rosenthal characterized the Banach spaces containing ℓ^1 in 1974 [Ros74], Bourgain, Fremlin and Talagrand studied pointwise compact sets of functions of the first Baire class in 1978 [BFT78], and Todorčević proved his trichotomy in 1999 [Tod99]; Argyros, Dodos and Kanellopoulos proved the heptachotomy in 2008, by the infinite-dimensional Ramsey theory of trees and perfect sets [ADK08]. The convexity theorem, Theorem 19.1, is a special case of a known result: as Todorčević pointed out to us, hereditarily separable locally connected Rosenthal compacta are metrizable, by his Theorems 5 and 7 [Tod99] and by Lemma 2.4 of Gruenhage [Gru88]. Kim and Pillay proved the independence theorem and the characterization of simplicity in 1997 [KP97], and Kim the symmetry of forking in simple theories in 1998 [Kim98]; Ben Yaacov extended simplicity to continuous logic [BY03]. The random graph was described by Ackermann in 1937 [Ack37]; Erdős and Rényi proved in 1963 that it is the almost sure countable random graph [ER63], and Rado gave his explicit construction of it in 1964 [Rad64]. Fraïssé introduced his limits in 1954 [Fra54], and Ben Yaacov extended them to metric structures [BY15]. Smoothly approximable structures were introduced by Lachlan; Kantor, Liebeck and Macpherson classified the primitive ones in 1989 [KLM89], and Cherlin and Hrushovski developed their structure theory [CH03]. The symplectic description of the Pauli group and of stabilizer codes is due to Gottesman [Got97] and to Calderbank, Rains, Shor and Sloane [CRSS98], and the five-qubit code to Laflamme, Miquel, Paz and Zurek [LMPZ96] and to Bennett, DiVincenzo, Smolin and Wootters [BDSW96]. Keisler introduced randomizations in 1999 [Kei99], and Ben Yaacov proved in 2013 that randomization destroys simplicity in the presence of the independence property [BY13]. Lee and Yang proved their theorem for the Ising ferromagnet in 1952 [LY52]. Asano [Asa70] and Suzuki and Fisher [SF71] extended it to quantum spin systems, including anisotropic Heisenberg ferromagnets in transverse fields, and Fröhlich and Rodriguez gave a general formulation [FR12, Theorem 5]. The approximation of a quantum spin system by a classical Ising system in one more dimension, which Variatio 18 uses to derive the zero-free half-plane $\operatorname{Re} h > 0$ directly, is due to Suzuki [Suz76]. Hong, Ou and Mandel observed their two-photon interference in 1987 [HOM87]. Model theory has met operator algebras and quantum theory along other routes: Farah, Hart and Sherman studied stability in the model theory of operator algebras [FHS13], where it is a property of the theory of an algebra rather than of a single observable; Zilber gave a model-theoretic semantics of the canonical commutation relation

[Zil16]; Ben Yaacov showed that randomization preserves NIP [BY09], a result of which the uniform NIP of comparisons of physical quantities is an elementary quantitative analogue; and Hrushovski and Pillay studied invariant measures in NIP theories [HP11].

The study of Shelah's dividing lines for bipartite observables grew out of conversations of the author with Eduardo Dueñez, and the results on observables presented here come from that study, circulated in the preprints [DI26a, DI26b, DI26c]; Table 9 lists them with their relation to earlier work.

Acknowledgments

I thank Eduardo Dueñez for the conversations out of which these results grew, and for his permission to present them here. I thank Stevo Todorčević for pointing out to me that Theorem 19.1 is a special case of a known result, and for the references that establish it.

The form of this memoir is that of J. S. Bach's Goldberg Variations, published in 1741: an aria, thirty variations in which every third variation up to the twenty-seventh is a canon, at intervals rising from the unison to the ninth, a quodlibet as the thirtieth, and the aria again. The section titles follow Bach's markings. In Figure 13 I measure this memoir against the music: the share of each variation in six performances of the Goldberg Variations, four on the piano and two on the harpsichord, and its share of the words here.

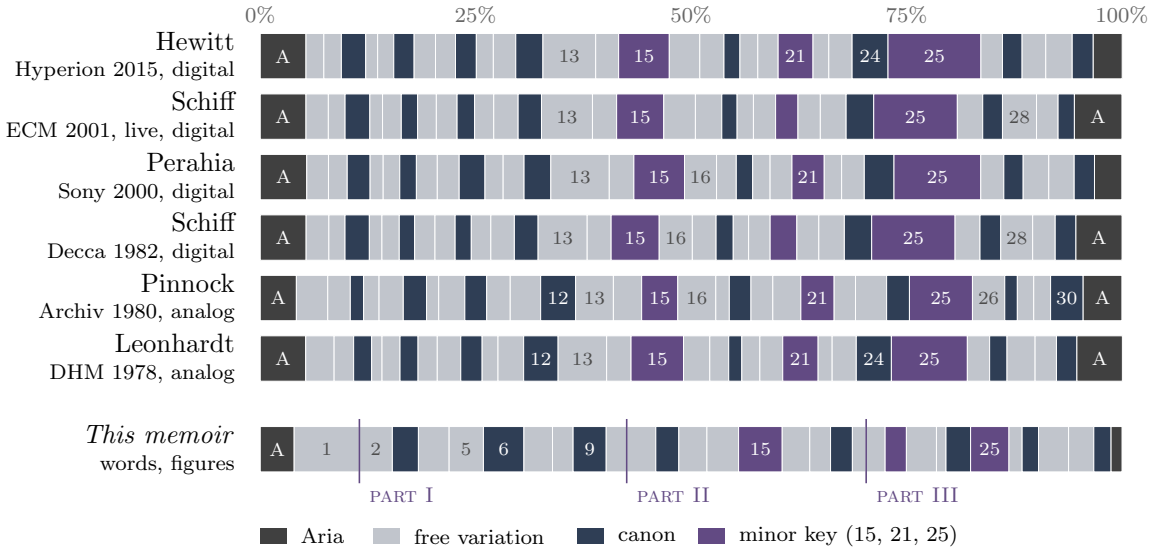


FIGURE 13. The proportions of the Goldberg Variations in six performances and in this memoir.

Contributions and antecedents

In Table 9 we distinguish the new results and constructions of this memoir from its applications and readings of classical theorems, and we record the antecedents of each. The memoir develops and unifies work circulated jointly with Eduardo Dueñez in three preprints [**DI26a**, **DI26b**, **DI26c**], together with the further constructions and interpretations recorded in the table.

Table 9: Contributions of the memoir.

Contribution	Location	Content	Relation to earlier work
Part I. Stability			
Stability as the agreement of two orders of idealization	Theorem 3.1 and Theorem 5.1	A bipartite observable is stable exactly when the two products of states, one for each order of idealization, agree at every pair of states; seven equivalent forms, among them the simulation of idealized preparations by finite mixtures and the agreement of iterated integrals.	Grothendieck, Arens, Davis–Figiel–Johnson–Pełczyński, Mazur, Pták–Simons and Harrington, read here for observables
The independent product	Theorem 6.4	On the stable observables the two orders of idealization define a state, symmetric in the two systems, which is the only separately weak* continuous extension of the product of actual states.	New formulation on the operator system of stable observables; Arens regularity gives the extension, and Harrington's lemma is the two-valued counterpart
The quantum-logical meet	Proposition 6.3	Two uniformly stable projections in the minimal tensor product can have an unstable meet; stability passes to norm limits but not to strong limits.	New counterexample for the meet; the passage to norm limits follows from the classical theory of weak compactness

Contribution	Location	Content	Relation to earlier work
Uniform stability, discrimination and learning	Theorem 7.1 and Example 7.2	Uniform stability is equivalent to bounded tree height, uniform in the threshold, to stability in an ultrapower, to polynomial growth of adaptive transcripts and to a finite online mistake bound; a stable observable need not be uniformly stable.	Shelah, Sauer–Shelah and Littlestone, applied to observables (compare Chase and Freitag); the sector example is new
Bounds independent of the dimension	Theorem 8.1	The trees of the swap test and of the Hong–Ou–Mandel coincidence at margin δ have height at most $4/\delta^2$ in every dimension.	Novikoff's perceptron bound, applied here
Simulation by finite mixtures	Theorem 9.1	Under stability, every idealized preparation is simulated uniformly over all probes by finite mixtures from any tail of any approximating net; for the overlap measurements an equal-weight mixture of $N \geq 1/\varepsilon^2$ terms suffices in every dimension, and the order of idealization cannot be dropped.	New characterization for observables, from weak compactness and Mazur's theorem; the quantitative bounds use Maurey's empirical method
Randomization	Theorem 10.1	No countably additive randomization detects instability; finitely additive means detect it exactly.	Pták–Simons, read here for observables; the countably additive case follows from the density-operator barycenter
Part II. NIP			
NIP as the first Baire class	Theorem 11.1	On the state space of the probe algebra, stability is continuity of every idealized response, NIP is membership of every idealized response in the first Baire class, and failure of NIP gives a copy of $\beta\mathbb{N}$.	Rosenthal, Bourgain–Fremlin–Talagrand and Grothendieck, applied to the probe algebra introduced here
Three classes for observables	Theorem 19.1 and Theorem P.32	Mixtures leave exactly three topological classes for the response compacta of NIP observables; every trace-norm dense sequence of preparations contains a tree subfamily equivalent to D_6 or D_7 in the first countable class that is not hereditarily separable, and to D_5 in the class that is not first countable.	New classification for observables, from the metrizability criterion of Todorčević and Gruenhage and the canonical embeddings of Argyros, Dodos and Kanellopoulos; a new short proof is given for convex compacta

Contribution	Location	Content	Relation to earlier work
Physical orders in every class	Theorem 16.1 and Corollary P.36	Photon numbers are NIP$^{\mathrm{met}}$; positions and quadratures are NIP$^{\mathrm{fc}}$ but not NIP$^{\mathrm{hs}}$, with idealized responses forming a subspace of the Helly space; the causal order and the phases are NIP but not NIP$^{\mathrm{fc}}$; lacunary fringes are not NIP; comparisons are uniformly NIP with bounds independent of the Hilbert spaces.	New identifications of response compacta, by classical order and function-space methods
The heptachotomy in physics	Theorem 15.1	Each of the seven minimal families of Argyros, Dodos and Kanellopoulos is realized by a tree of pure preparations of photons, a particle in a box or a spin chain.	New
Spontaneous magnetization as a split pair	Theorem 18.1, Theorem P.54	The Ising magnetization with offsets is stable exactly when the spontaneous magnetization vanishes, and is otherwise NIP$^{\mathrm{fc}}$ but not NIP$^{\mathrm{hs}}$, its two pure phases forming split pairs; the same holds in a transverse field, whose partition function has no zeros off the imaginary axis.	New classification; the thermodynamic input, including the Lee–Yang theorem in a transverse field, is classical
Asymmetry of NIP	Corollary P.52	The scanning observable is NIP$^{\mathrm{met}}$ for Alice's responses, is not uniformly NIP, and is not NIP for Bob's responses.	Physical realization, with mixed preparations, of an example of Khanaki and Pillay [**KP18**]; the metrizable response compactum is identified here
Measures on the split interval	Remark P.30	The space of probability measures on the split interval is not hereditarily separable.	Corrects a remark of Godefroy
Part III. Simplicity			
Forking and logical errors	Theorem 25.1 and Corollary 25.2	An undetectable error with nonzero label is harmless exactly when the type of its label over the stabilizer labels forks over the empty set; for a code that stores at least one qubit, the distance is the least weight of an undetectable error with nonzero label whose type does not fork.	New coding interpretation of the forking criterion for symplectic spaces and of the stabilizer criterion for harmless errors

Contribution	Location	Content	Relation to earlier work
Homogeneity as Clifford circuits	Theorem 24.2 and Corollary 24.3	Every linear isometry between finite-dimensional spaces of Pauli labels is implemented, modulo phases, by a Clifford circuit, and every stabilizer code has a Clifford encoding circuit.	Witt's theorem and the classical Clifford realization of symplectic maps and stabilizer encodings, read here as the homogeneity of a Fraïssé limit
The gluing of syndromes	Proposition 26.1	Two syndromes prescribed for checks designed independently are produced by one error when they agree on the common checks.	Elementary symplectic linear algebra, read here as the independence theorem of Kim and Pillay
Graph states realize every relation	Proposition 27.1	Every relation between two finite sets is a pattern of anticommutation between single-qubit flips and the stabilizer generators of a graph state.	Standard graph-state stabilizers, used here to connect finite relations with Pauli commutation
Mixing destroys simplicity	Theorem 28.1	For two-valued diagonal observables, mixtures turn the independence property of the pure responses into the tree property of the second kind of a fixed affine formula, witnessed in an elementary extension; a simple diagonal observable is NIP.	New for the mixed response structure; analogous to, but distinct from, Ben Yaacov's theorem on randomizations
Actual and idealized witnesses	Theorem 29.1 and Propositions P.69 and P.66	For the Pauli observable, actual coset mixtures witness the tree property of the second kind for the response predicate, with the bound 2/3, sharp for this array; for the Rado observable no fixed quantifier-free formula in one probe variable has a witnessing array whose preparation parameters are all actual. The probes of the Pauli observable are the Pauli channels on countably many qubits.	New witnesses and a new obstruction; the Fourier description of Pauli channels is standard

Contribution	Location	Content	Relation to earlier work
One relation through three lines	Proposition 30.1	The commutation phase organizes three regimes: stable fringes for bounded wavenumbers, fringes that are not NIP for the wavenumbers 8^n, and the simple binary Pauli structure, whose mixed response structure has the tree property of the second kind.	Synthesis of the preceding results through the Weyl relation and its binary counterpart

Bibliography

[ABDCBH97] Noga Alon, Shai Ben-David, Nicolò Cesa-Bianchi, and David Haussler, *Scale-sensitive dimensions, uniform convergence, and learnability*, Journal of the ACM **44** (1997), no. 4, 615–631.

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

[ADK08] Spiros A. Argyros, Pandelis Dodos, and Vassilis Kanellopoulos, *A classification of separable Rosenthal compacta and its applications*, Dissertationes Mathematicae **449** (2008), 1–52.

[Are51] R. Arens, *The adjoint of a bilinear operation*, Proc. Amer. Math. Soc. **2** (1951), 839–848.

[Art57] E. Artin, *Geometric Algebra*, Interscience, New York, 1957.

[Asa70] Taro Asano, *Lee–Yang theorem and the Griffiths inequality for the anisotropic Heisenberg ferromagnet*, Physical Review Letters **24** (1970), 1409–1411.

[Ass83] Patrick Assouad, *Densité et dimension*, Annales de l'Institut Fourier **33** (1983), no. 3, 233–282.

[BCWdW01] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, *Quantum fingerprinting*, Phys. Rev. Lett. **87** (2001), 167902.

[BDPSS09] S. Ben-David, D. Pál, and S. Shalev-Shwartz, *Agnostic online learning*, Proc. 22nd Conference on Learning Theory (COLT 2009), 2009.

[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.

[BFT78] J. Bourgain, D. H. Fremlin, and M. Talagrand, *Pointwise compact sets of Baire-measurable functions*, Amer. J. Math. **100** (1978), no. 4, 845–886. MR 509077

[BG09] Jakob E. Björnberg and Geoffrey R. Grimmett, *The phase transition of the quantum Ising model is sharp*, Journal of Statistical Physics **136** (2009), 231–273, arXiv:0901.0328.

[BK90] Christian Borgs and Roman Kotecký, *A rigorous theory of finite-size scaling at first-order phase transitions*, Journal of Statistical Physics **61** (1990), 79–119.

[BLW96] Peter L. Bartlett, Philip M. Long, and Robert C. Williamson, *Fat-shattering and the learnability of real-valued functions*, Journal of Computer and System Sciences **52** (1996), no. 3, 434–452.

[BY03] I. Ben Yaacov, *Simplicity in compact abstract theories*, J. Math. Log. **3** (2003), 163–191.

[BY09] ———, *Continuous and random Vapnik–Chervonenkis classes*, Israel J. Math. **173** (2009), 309–333.

[BY13] Itaï Ben Yaacov, *On theories of random variables*, Isr. J. Math. **194** (2013), 957–1012.

[BY14] ———, *Model theoretic stability and definability of types, after A. Grothendieck*, Bull. Symb. Log. **20** (2014), no. 4, 491–496. MR 3294276

[BY15] I. Ben Yaacov, *Fraïssé limits of metric structures*, J. Symbolic Logic **80** (2015), 100–115.

[BYBHU08] Itaï Ben Yaacov, Alexander Berenstein, C. Ward Henson, and Alexander Usvyatsov, *Model theory for metric structures*, Model theory with applications to algebra and analysis. Vol. 2, London Math. Soc. Lecture Note Ser., vol. 350, Cambridge Univ. Press, Cambridge, 2008, pp. 315–427. MR 2436146

[BYU10] Itaï Ben Yaacov and Alexander Usvyatsov, *Continuous first order logic and local stability*, Trans. Amer. Math. Soc. **362** (2010), no. 10, 5213–5259. MR 2657678

[CF19] H. Chase and J. Freitag, *Model theory and machine learning*, Bull. Symbolic Logic **25** (2019), 319–332.

[CH03] G. Cherlin and E. Hrushovski, *Finite Structures with Few Types*, Annals of Mathematics Studies, vol. 152, Princeton University Press, 2003.

[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.

[CS16] A. Chernikov and S. Shelah, *On the number of Dedekind cuts and two-cardinal models of dependent theories*, J. Inst. Math. Jussieu **15** (2016), 771–784.

[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.

[Deb13] Gabriel Debs, *Descriptive aspects of Rosenthal compacta*, Recent Progress in General Topology III, Atlantis Press, 2013, pp. 205–227.

[Deu01] F. Deutsch, *Best Approximation in Inner Product Spaces*, CMS Books in Mathematics, Springer, New York, 2001.

[DFJP74] W. J. Davis, T. Figiel, W. B. Johnson, and A. Pełczyński, *Factoring weakly compact operators*, J. Functional Analysis **17** (1974), 311–327.

[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 Todorčević's trichotomy*, preprint; arXiv:2610.12908, 2026.

[DI26c] ———, *Classification of bipartite observables III: simplicity, mixtures and quantum error correction*, preprint; arXiv:2610.22205, 2026.

[DIMW+26] E. Dueñez, J. Iovino, T. Matos-Wiederhold, L. Salvetti, and F. D. Tall, *Complexity of deep computations via topology of function spaces*, 2026, arXiv:2601.00528.

[Ebe47] W. F. Eberlein, *Weak compactness in Banach spaces. I*, Proc. Nat. Acad. Sci. U.S.A. **33** (1947), 51–53.

[Eng89] Ryszard Engelking, *General topology*, second ed., Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989, Translated from the Polish by the author. MR 1039321

[ER63] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta Math. Acad. Sci. Hungar. **14** (1963), 295–315.

[FHS13] Ilijas Farah, Bradd Hart, and David Sherman, *Model theory of operator algebras I: Stability*, Bull. Lond. Math. Soc. **45** (2013), no. 4, 825–838.

[FR12] Jürg Fröhlich and Pierre-François Rodriguez, *Some applications of the Lee–Yang theorem*, Journal of Mathematical Physics **53** (2012), 095218, arXiv:1205.6643.

[Fra54] R. Fraïssé, *Sur l'extension aux relations de quelques propriétés des ordres*, Ann. Sci. École Norm. Sup. **71** (1954), 363–388.

[FV17] Sacha Friedli and Yvan Velenik, *Statistical mechanics of lattice systems: A concrete mathematical introduction*, Cambridge University Press, 2017.

[Gan40] V. Gantmacher, *Über schwache totalstetige Operatoren*, Mat. Sbornik N.S. **7(49)** (1940), 301–308.

[Gin69] Jean Ginibre, *Existence of phase transitions for quantum lattice systems*, Communications in Mathematical Physics **14** (1969), 205–234.

[GM22] Eli Glasner and Michael Megrelishvili, *Todorčević' trichotomy and a hierarchy in the class of tame dynamical systems*, Transactions of the American Mathematical Society **375** (2022), no. 7, 4513–4548.

[God80] Gilles Godefroy, *Compacts de Rosenthal*, Pacific J. Math. **91** (1980), no. 2, 293–306. MR 615679

[Got97] D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. thesis, California Institute of Technology, 1997, arXiv:quant-ph/9705052.

[Gri64] Robert B. Griffiths, *Peierls proof of spontaneous magnetization in a two-dimensional Ising ferromagnet*, Physical Review **136** (1964), A437–A439.

[Gro52] A. Grothendieck, *Critères de compacité dans les espaces fonctionnels généraux*, Amer. J. Math. **74** (1952), 168–186. MR 47313

[Gro84] Ulrich Groh, *Uniform ergodic theorems for identity preserving Schwarz maps on W^*-algebras*, J. Operator Theory **11** (1984), no. 2, 395–404. MR 86a:46079

[Gru88] Gary Gruenhage, *On the existence of metrizable or Sorgenfrey subspaces*, General Topology and its Relations to Modern Analysis and Algebra VI (Proc. Sixth Prague Topological Symposium, 1986) (Zdeněk Frolík, ed.), Research and Exposition in Mathematics, vol. 16, Heldermann, Berlin, 1988, pp. 223–230.

[Hei80] S. Heinrich, *Ultraproducts in Banach space theory*, J. Reine Angew. Math. **313** (1980), 72–104.

[HH84] V. Harnik and L. Harrington, *Fundamentals of forking*, Ann. Pure Appl. Logic **26** (1984), 245–286.

[HOM87] C. K. Hong, Z. Y. Ou, and L. Mandel, *Measurement of subpicosecond time intervals between two photons by interference*, Phys. Rev. Lett. **59** (1987), 2044–2046.

[HP11] E. Hrushovski and A. Pillay, *On NIP and invariant measures*, J. Eur. Math. Soc. **13** (2011), 1005–1061.

[Iov99a] José N. Iovino, *Stable Banach spaces and Banach space structures. I. Fundamentals*, Models, algebras, and proofs (Bogotá, 1995), Dekker, New York, 1999, pp. 77–95. MR 2000h:03064a

[Iov99b] ———, *Stable models and reflexive Banach spaces*, J. Symbolic Logic **64** (1999), no. 4, 1595–1600. MR 2001e:03068

[Kad51] Richard V. Kadison, *A representation theory for commutative topological algebra*, Memoirs of the American Mathematical Society, vol. 7, American Mathematical Society, 1951.

[Kei87] H. J. Keisler, *Measures and forking*, Ann. Pure Appl. Logic **34** (1987), 119–169.

[Kei99] H. Jerome Keisler, *Randomizing a model*, Adv. Math. **143** (1999), no. 1, 124–158. MR 1680650

[Kei23] H. Jerome Keisler, *Model theory for real-valued structures*, Beyond First-Order Model Theory, Vol. 2 (José Iovino, ed.), CRC Press, 2023, arXiv:2005.11851, pp. 43–101.

[Kim98] B. Kim, *Forking in simple unstable theories*, J. London Math. Soc. (2) **57** (1998), 257–267.

[KLM89] W. M. Kantor, M. W. Liebeck, and H. D. Macpherson, $\aleph_0$*-categorical structures smoothly approximated by finite substructures*, Proc. London Math. Soc. (3) **59** (1989), no. 3, 439–463.

[KM81] Jean-Louis Krivine and Bernard Maurey, *Espaces de Banach stables*, Israel J. Math. **39** (1981), no. 4, 273–295. MR 83a:46030

[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 1600895

[KP18] Karim Khanaki and Anand Pillay, *Remarks on the NIP in a model*, Math. Log. Q. **64** (2018), no. 6, 429–434, arXiv:1706.04674.

[KR83] Richard V. Kadison and John R. Ringrose, *Fundamentals of the theory of operator algebras, Vol. I*, Academic Press, 1983.

[Lit88] N. Littlestone, *Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm*, Machine Learning **2** (1988), 285–318.

[LMPZ96] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, *Perfect quantum error correcting code*, Phys. Rev. Lett. **77** (1996), 198–201.

[LS81] Elliott H. Lieb and Alan D. Sokal, *A general Lee–Yang theorem for one-component and multicomponent ferromagnets*, Communications in Mathematical Physics **80** (1981), no. 2, 153–179.

[LY52] T. D. Lee and C. N. Yang, *Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model*, Physical Review **87** (1952), 410–419.

[Maz33] S. Mazur, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1933), 70–84.

[New74] Charles M. Newman, *Zeros of the partition function for generalized Ising systems*, Communications on Pure and Applied Mathematics **27** (1974), 143–159.

[Nov62] A. B. J. Novikoff, *On convergence proofs on perceptrons*, Proc. Symposium on the Mathematical Theory of Automata, vol. 12, Polytechnic Institute of Brooklyn, 1962, pp. 615–622.

[Pei36] Rudolf Peierls, *On Ising's model of ferromagnetism*, Proceedings of the Cambridge Philosophical Society **32** (1936), 477–481.

[Pis81] G. Pisier, *Remarques sur un résultat non publié de B. Maurey*, Séminaire d'Analyse Fonctionnelle 1980–1981, Exp. V, École Polytechnique, Palaiseau, 1981.

[Ptá63] V. Pták, *A combinatorial lemma on the existence of convex means and its application to weak compactness*, Convexity, Proc. Sympos. Pure Math., vol. 7, Amer. Math. Soc., 1963, pp. 437–450.

[Rad64] R. Rado, *Universal graphs and universal functions*, Acta Arith. **9** (1964), 331–340.

[Ray83] Y. Raynaud, *Espaces de Banach superstables, distances stables et homéomorphismes uniformes*, Israel J. Math. **44** (1983), no. 1, 33–52. MR 84j:46024

[Ray02] Yves Raynaud, *On ultrapowers of non commutative L_p spaces*, J. Operator Theory **48** (2002), no. 1, 41–68.

[Rod22] José Rodríguez, *ε-weakly precompact sets in Banach spaces*, Studia Math. **262** (2022), no. 3, 327–360, arXiv:2102.01546.

[Ros74] H. P. Rosenthal, *A characterization of Banach spaces containing l^1*, Proc. Nat. Acad. Sci. U.S.A. **71** (1974), 2411–2413. MR 358307

[Row64] F. Rowbottom, *The Łoś conjecture for uncountable theories*, Notices of the Amer. Math. Soc. **11** (1964), 248–249.

[RST15] A. Rakhlin, K. Sridharan, and A. Tewari, *Online learning via sequential complexities*, J. Mach. Learn. Res. **16** (2015), 155–186.

[Sau72] N. Sauer, *On the density of families of sets*, J. Combinatorial Theory Ser. A **13** (1972), 145–147. MR 307902

[SF71] Masuo Suzuki and Michael E. Fisher, *Zeros of the partition function for the Heisenberg, ferroelectric, and general Ising models*, Journal of Mathematical Physics **12** (1971), 235–246.

[SG73] Barry Simon and Robert B. Griffiths, *The $(\phi^4)_2$ field theory as a classical Ising model*, Communications in Mathematical Physics **33** (1973), 145–164.

[She69] S. Shelah, *Stable theories*, Israel J. Math. **7** (1969), 187–202.

[She71] Saharon Shelah, *Stability, the f.c.p., and superstability; model theoretic properties of formulas in first order theory*, Ann. Math. Logic **3** (1971), no. 3, 271–362. MR 0317926

[She72] ———, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific J. Math. **41** (1972), 247–261. MR 307903

[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 1083551

[Sim68] S. Simons, *The iterated limit condition, a Fubini theorem and weak compactness*, Math. Ann. **176** (1968), 87–95.

[Sim93] Barry Simon, *The statistical mechanics of lattice gases*, vol. I, Princeton University Press, Princeton, NJ, 1993.

[Sim15a] Pierre Simon, *A guide to NIP theories*, Lecture Notes in Logic, vol. 44, Association for Symbolic Logic, Chicago, IL; Cambridge University Press, Cambridge, 2015. MR 3560428

[Sim15b] ———, *Rosenthal compacta and NIP formulas*, Fund. Math. **231** (2015), no. 1, 81–92. MR 3361236

[Suz76] Masuo Suzuki, *Relationship between d-dimensional quantal spin systems and $(d+1)$-dimensional Ising systems: equivalence, critical exponents and systematic approximants of the partition function and spin correlations*, Progress of Theoretical Physics **56** (1976), 1454–1469.

[Tit39] E. C. Titchmarsh, *The theory of functions*, second ed., Oxford University Press, Oxford, 1939.

[Tod97] Stevo Todorcevic, *Topics in topology*, Lecture Notes in Mathematics, vol. 1652, Springer Berlin, Heidelberg, 1997.

[Tod99] Stevo Todorčević, *Compact subsets of the first Baire class*, J. Amer. Math. Soc. **12** (1999), no. 4, 1179–1212. MR 1685782

[Š40] V. Šmulian, *Über lineare topologische Räume*, Mat. Sbornik N.S. **7(49)** (1940), 425–448.

[Zil16] B. Zilber, *The semantics of the canonical commutation relation*, arXiv:1604.07745, 2016.