The price of precision: usable entanglement in nonlocal games and the geometry of groups

Seth Douglas Nidhal Mghirbi

Version 1.1.0, October 2026

Abstract

In a nonlocal game of value one, entanglement demand at loss ε is the infimum of the one-way distillable entanglement of an original strategy state losing at most ε. For synchronous games with tested diagonals and constraint-variable consistency games, this demand diverges exactly when no finite-dimensional strategy is perfect. Every good enough mixed strategy yields any fixed number of Bell pairs at any fixed fidelity below one by local isometries. Under a promise, a sufficient loss is computable for rational synchronous or binary constraint games, but bounded demand is undecidable among games of value one. Agreement games have the same demand in both directions up to a factor three in loss. A sign relation and an even dilation, mirrored by both players, force a logarithmic lower bound. Theorem 1.1 of our companion "The uncomputable price of precision" (10.5281/zenodo.23736188) realizes every positive Δ_2^0 logarithmic coefficient with a fixed game on the original mixed state. We classify the positive pure original-state entropy-coefficient spectrum as the positive Δ_2^0 reals; the full usable spectrum remains open. The least usable entanglement compatible with M approximate Pauli pairs at pairwise error $\kappa \leq 1/100$ is of order $\min\{M, \kappa^{-2}\}$. Group geometry gives logarithmic laws, an almost inverse-loss law, and a point-Pauli game with demand between $\varepsilon^{-\beta}$ and $\varepsilon^{-1/4}$, where $\beta = 0.0734595360384268649\ldots$. Its lower bound uses explicit far-commutator fillings; a magnetic-cycle estimate obstructs improving the fixed-holonomy subdivision alone. Under exact group relations, demand counts qubits addressable by short words, with polynomial and exponential laws. Conditional on attainment and tail bounds, the I_{3322} optimum needs bounded usable entanglement.

The authors developed this work with AI models from more than one developer and take responsibility for every claim. How the work was done, and what the machine checks cover, is described at the end of the paper.

Contents

1 Introduction **2**

2 Preliminaries **13**

I. The engine **20**

3 The qubit count in an arbitrary state **20**

4 Transfer and return with the state kept **21**

5 Compiling groups into games **26**

6 Operational extraction on the original state **34**

II. The law 45

7 The law: unreachable perfection forces usable entanglement 45

8 Directional invariance 53

III. The spectrum 57

9 Stable spectral transport and the Slofstra–Vidick game 57

10 Forced branching: the lamp sources 62

11 Games with a dense set of coefficients 69

12 The configuration-lamp game: usable entanglement 88

13 The point-Pauli game: a polynomial window 91

14 Exact sources: the price of an address 99

IV. Limits 104

15 Obstructions and refutations 104

16 A Bell-gap contrast and finite blocks 112

17 The logarithmic floor: a class theorem and a criterion 113

18 Open problems 118

19 Discussion 121

How this work was done 122

A Proof of the white-box protocol 123

B Status of the results 125

C Reproducibility and computational scope 128

D Results used from our companion papers 130

E The Ji–Leung–Vidick game with a tapered reservoir 131

F The floor: candidate classes and method obstructions 133

G The point window: fillings and subdivision obstructions 142

1 Introduction

1.1 The question

In a nonlocal game a referee asks two players questions and accepts if their answers satisfy a fixed predicate. The players cannot communicate, but they may share an entangled state. Some games of quantum value one cannot be won exactly by any finite-dimensional strategy, although finite strategies come arbitrarily close [132]. Such nonattainment forces the dimension of finite near-winning strategies to diverge as their loss ε tends to zero. A

large dimension certifies only a large space, which may hold a separable state, so we ask when nonattainment also forces the distillable entanglement of their original mixed states to diverge, and at what rate.

"Entanglement" can mean several different things here, and they behave differently. The local *dimension* of a strategy, or the Schmidt rank of a pure one, counts the size of the Hilbert space. The *entanglement entropy* of a pure strategy counts ebits. The *entanglement of formation* of a mixed strategy is the least average entropy over pure decompositions. Dimension, marginal entropy and entanglement of formation do not in general lower-bound the distillable entanglement of an arbitrary mixed strategy: a large dimension or a large marginal entropy is compatible with a separable state, and a pure decomposition is a mathematical expansion, not a label the players can read. A mixed state can even have positive entanglement of formation and no distillable entanglement at all [72]. For pure states, entropy does equal the asymptotic Bell-pair rate [12]. We seek guarantees on the actual mixed state without access to a label of one of its pure decompositions, so the resource we bound is the *distillable* entanglement of the actual shared state, the rate at which local operations and classical communication turn copies of it into Bell pairs. We work with the one-way version, $E_{\mathrm{D}}^{A\to B}$, and define the *demand* of a game as

$$d_G^{\rightarrow}(\varepsilon) = \inf\{E_{\mathrm{D}}^{A\to B}(\rho_{AB}) : \text{ some finite-dimensional strategy on } \rho_{AB} \text{ has loss} \le \varepsilon\}.$$

The infimum is over arbitrary mixed states and measurements, so a lower bound on $d_G^{\rightarrow}$ is a statement about every near-winning strategy, with all its noise and all its correlations with an environment. Throughout, the *original state* is the shared state ρ_{AB} that the players hold before they measure. Playing the game consumes it, and we make no claim about entanglement that survives play. That question is addressed by Arnon-Friedman and Bancal, who certify, device-independently, one-shot distillable entanglement in the state that remains after a test [3], and Bhavsar, Wooltorton and Bae give a composable device-independent certification of states that a source prepares as a finite sequence of independent, not necessarily identical, states [19]. Our bounds concern the state before the game is played, and they hold for every strategy of a fixed game rather than for a protocol that tests a source.

Winning with loss ε controls the players' operators to precision about $\sqrt{\varepsilon}$, measured in their shared state. Through this precision a game sees a group or an algebra, its *source*, and the demand is a property of that source at precision $\sqrt{\varepsilon}$. A source group always comes with marked data: a central involution J, which must act as -1 on the state, or at least stay visibly away from 1 there. Without this requirement the trivial representation would satisfy every relation at no cost. Our thesis is that *the geometry of groups is the price of precision*: the absence of finite-dimensional representations makes usable entanglement diverge, and quantitative invariants of the source (areas of fillings, spectra of dilations, the growth of short addresses) set how fast.

1.2 What was known

Mančinska and Vidick gave a two-player game, with answers of unbounded length, whose optimal value no finite amount of entanglement attains [96]; the games below have finite answer sets. Slofstra showed that every finitely presented group embeds in a solution group [133], and used embeddings of this kind to prove that the set of quantum correlations is not closed and that deciding whether a linear system game has a finite-dimensional perfect strategy is undecidable [132]. Non-closure was later proved by other routes: via graphs [59], via correlation matrices and factorizable channels that require infinite-dimensional ancillas [106], and via embezzlement [36]; and Beigi separated the quantum, spatial quantum and approximate quantum correlation sets [11]. Slofstra and Vidick connected the entanglement of a linear

system game to the hyperlinear profile of its solution group, and built a fixed game whose least local dimension lies between $\varepsilon^{-1/6}$ and $\varepsilon^{-1/2}$, from a group whose profile lies between $\varepsilon^{-2/3}$ and ε^{-1} [134, Theorem 1.1 and Proposition 3.2]. Slofstra then built a finitely presented group with at least subexponential hyperlinear profile, and from it a two-player game of commuting-operator value one in which success probability $1-\varepsilon$ needs dimension at least $\exp(c\,\varepsilon^{-a})$ for some $a > 0$; whether that game has quantum value one is left open there [131, Theorem 1.2 and Corollary 1.4]. Ji, Leung and Vidick gave a three-player game of quantum value one in which success probability $1-\varepsilon$ needs local dimension $2^{\Omega(\varepsilon^{-c})}$ [74, Theorem 1]. Coladangelo gave a two-player game, based on embezzlement, in which a strategy whose expected score is within ε of the optimum needs dimension $2^{\Omega(\varepsilon^{-1/8})}$ [36, Theorems 1 and 5]. His game is scored by a real-valued function, so his ε is a gap below an optimal Bell score, not a loss from winning probability one; his exponent and the exponents below are therefore measured under different contracts and are not directly comparable. Our companion paper gives an explicit linear system game of quantum value one whose near-winners need entanglement entropy, and mixed strategies entanglement of formation, $\varepsilon^{-1+o(1)}$, so that $\log \mathcal{E}(G,\varepsilon) = \varepsilon^{-1+o(1)}$ [101, Theorem A]; it is built on the configuration lamps of our earlier paper [54]. Ostrev and Vidick showed that the CHSH(n) games need $\Omega(\varepsilon^{-1/5})$ ebits at relative precision ε, that is, at a bias within a factor $1-\varepsilon$ of the optimum [111], and the companion paper improves this to the optimal order $\Theta(\min(n, 1/\varepsilon))$ for entanglement entropy and formation [101, Theorem E]. None of these bounds the distillable entanglement of a mixed near-winner.

Near-perfect strategies can often be rounded to better ones, at the price of replacing the state. Paddock rounds near-perfect strategies of constraint system and synchronous games to strategies on a new maximally entangled state, close to the original measurements on its support [112, Corollaries 4.14 and 4.28]. Vidick approximates almost synchronous strategies by convex combinations of projective strategies on maximally entangled states [145, Theorem 3.1], and Marrakchi and de la Salle extend this to infinite-dimensional strategies [97]; for projection games Culf rounds to approximately tracial strategies [45]. None of these extracts Bell pairs from the original state. Robust self-testing of many EPR pairs [107, 29, 38, 98] and constant-sized self-tests of unbounded dimension [63, 94] are established, and, under mild conditions on the ideal strategy, self-tests proved for pure states with projective measurements extend to arbitrary mixed states and POVMs [8, Theorem 4.1]. For synchronous games, robust self-testing against projective maximally entangled strategies lifts to all POVM strategies [143, Theorem 1.2], and, for extreme synchronous correlations, to arbitrary mixed-state strategies [31, Theorem 3.3]. For games with a robust game algebra, among them synchronous and binary constraint system games, Zhao characterizes robust self-testing through uniqueness of an amenable tracial state on the universal measurement algebra satisfying the game relations, with that unique trace finite-dimensional [149, Theorem 7.10]; Kar gives a corresponding formulation for synchronous game algebras [80, Theorem 5.5]. The results just cited certify a prescribed ideal strategy of a fixed finite dimension. Self-tests whose ideal has no finite-dimensional realization also exist, in the abstract-state formulation of self-testing of Paddock, Slofstra, Zhao and Zhou [115, Section 7]: a correlation that is both a commuting-operator and a quantum-approximate self-test [31, Section 6.1, Proposition 6.2], and a commuting-operator self-test for exact embezzlement of entanglement [113, Theorem 1.1]. Each asserts that one exact correlation is realized by a unique abstract state. Theorem A below is a statement of another kind. Its games have no finite-dimensional perfect strategy, it specifies no ideal strategy, and its conclusion is extraction of any prescribed number of Bell pairs, by one pair of local isometries, from the original mixed state of every sufficiently good finite-dimensional strategy of the fixed game. Universal distillation from unknown mixed states in a fixed finite dimension is also established [135, 138], building on protocols for compound sources [20].

In infinite dimensions the mechanism behind Theorem A is exact. For hyperfinite factors in Haag duality that generate all bounded operators, Keyl, Matsui, Schlingemann and Werner prove the infinite-one-copy-entanglement statement for every normal joint state when the factors are not of type I [81, Theorem 4.9]. Van Luijk, Stottmeister, Werner and Wilming remove hyperfiniteness for extraction from pure vector states by operations in the closure of LOCC [142, Theorem 54]; their all-density-matrix Bell statement retains approximate finite dimensionality. Theorem A uses the same dichotomy of types, but only for a tracial limit of finite-dimensional strategies, and it proves its finite-strategy mixed-state assertion directly, by extracting from those strategies themselves. Normal infinite-dimensional correlations that have no finite-dimensional realization were constructed by Coladangelo and Stark [39]; the finite-block conclusion for the classes of Theorem A (Remark 16.2) has a more restricted scope, perfect strategies of games in those classes.

Prior work includes operational distillation certification and quantitative entanglement bounds for specified self-tests. Our question is different: for a fixed finite consistency-tested game, does the original mixed-state one-way distillable entanglement of every near-winner diverge as its loss tends to zero, and can its growth be sharp?

1.3 The main results

We state the results in plain terms and point to the precise theorems. Logarithms are to base two. Two classes of games carry the law: *synchronous games* in which every diagonal question pair is tested with positive probability, and *constraint-variable consistency games*, in which Alice answers a satisfying assignment to a constraint, Bob the value of one of its variables, and every incidence is tested (Definition 7.1). Linear system games and Boolean constraint games belong to the second class. Here *Boolean* refers to the constraint variables: Bob answers one bit, but Alice answers an assignment to all variables of her constraint. Single-bit answers on both sides would not do, since a game in which both players answer one bit and which has a perfect finite-dimensional quantum strategy has a perfect classical one [32, Theorem 3].

Theorem A (Unreachable perfection forces usable entanglement; Theorem 7.2)**.** *Let G be a finite game of quantum value one in either class. Then G has no finite-dimensional perfect strategy if and only if $d_G^{\rightarrow}(\varepsilon) \to \infty$ as $\varepsilon \to 0$. More precisely, if G has no finite-dimensional perfect strategy, then for every r and $\gamma > 0$ there is $\varepsilon_0 > 0$ such that every finite-dimensional mixed strategy of loss at most ε_0 admits product local isometries producing r Bell pairs with fidelity at least $1 - \gamma$, the environment, untouched, being part of the junk.*

The proof takes a tracial limit of Bob's moments, uses the absence of finite perfect strategies to exclude every finite type I summand of the limit, approximates Pauli matrices there by bounded polynomials, and evaluates the polynomials back on the original strategies, where actual mirrors between the players' operators let them act on the original state. Three further results sharpen the picture. For rational games in class (S), and rational binary constraint-variable games in class (C), a threshold ε_0 for fixed r and rational γ is computable under the promise of quantum value one and no finite-dimensional perfect strategy, with finite rational certificates, once the polynomial moment relations are imposed on both sides (Theorem 7.5). Among games promised to have quantum value one, whether the demand is bounded or divergent is undecidable, and the divergent branch is not even semidecidable, already for games with contexts of arity four, sixteen answers for Alice and two for Bob (Theorem 7.6). The reduction keeps the promise by construction: each instance is gated against a fixed game of Slofstra that has value one and no finite-dimensional perfect strategy, so every instance has value one and only the existence of a finite perfect strategy varies. For the same rational games under the same promise, if the players have independent copies of the state and coherent white-box access to their own measurement circuits, as specified in

Theorem 7.7, one finite protocol, independent of the hidden state and its dimension, extracts the pairs from one retained copy; this model is not device-independent.

Theorem A gives no rate. A logarithmic floor does hold on an explicit class: if every near-winner of a fixed game supplies, with actual mirrors on its original state, unitaries satisfying a sign relation $WY = -YW$ and an even dilation $XYX^* = Y^m$, then $d_G^{\rightarrow}(\varepsilon) \geq \log(1/\varepsilon)/(4\log m) - O(\log\log(1/\varepsilon))$, and the Boolean constraint game compiled from Abels' finitely presented solvable group is an instance (Theorem 17.3 and Corollary 17.6). Whether every game of Theorem A has such a floor is open; Section 17 gives a sufficient criterion for a counterexample, and Appendix F the candidate classes that cannot meet it (Problem 18.2).

The demand has a direction, and one might fear that a game could demand usable entanglement only from Alice to Bob. It cannot, for the large class of agreement games, in which acceptance means agreement of labels of the two answers.

Theorem B (Directions are interchangeable at a loss factor of three; Theorems 8.2 and 8.3)**.** *For every finite agreement game, one channel applied by Alice alone, followed by a measurement of whole answers, turns a strategy of loss ε into one of loss at most 3ε whose two directional rates are equal and no larger than either original rate. So $d_G^{\rightarrow}(\varepsilon) \geq d_G^{\leftarrow}(3\varepsilon)$ and conversely, uniformly in the game. The factor three is optimal for this mechanism: no deterministic Alice-local channel that symmetrizes the state while keeping Bob's system, marginal and measurements fixed gives a smaller factor, whatever legal decoding follows.*

The optimality concerns this mechanism, not every comparison of the two demand profiles (Problem 18.6). For $\varepsilon \leq \frac{1}{2}$ the loss bound improves to $1-(1-\varepsilon)^3$ (Theorem 8.2), and the guarantee of this channel and decoder holds for a predicate exactly when the predicate has the agreement form (Theorem 8.3). The channel keeps Bob's operators and marginal, so every source budget measured on Bob is unchanged. Its output is the reflected state of Dutta and Faulkner [58], reached by a Petz recovery map [120] in the pretty-good form of Barnum and Knill [9, 17], and the factor three is the polynomial inequality $h^3 \geq 3h-2$ on $[-1,1]$ applied to KMS-adjoint correlation maps, related to the correlation-operator framework of Beigi's quantum maximal correlation [10].

Changing the finite presentation of the marked group changes the precision of every model by at most a constant factor and keeps the state (Proposition 5.3 and the paragraph after it). A constant factor changes $\alpha\log(1/e)$ only by $O(1)$ and leaves every exponent unchanged, so all the laws below are independent of the presentation; a change of precision by a power would change them, and none occurs. For the normalized trace, Chapman, Vidick and Yuen transfer stability moduli between presentations with the same kind of constant rescaling [30, Lemma 2.6]; here the state is kept, and conjugations are paid for through state commutators. For permutation and metric models, Cornulier's sofic profile is defined from the finite pieces of the group itself [42], and the profiles of Arzhantseva and Cherix do not depend on the generating set up to the usual equivalence [4]. The next theorem collects the fixed games whose demand we can compute; all are finite games fixed before ε is chosen.

Theorem C (A spectrum of fixed games; Theorems 9.9, 9.10, 11.1, 11.7, 11.9, 12.4 and 13.7 and Corollaries 11.6 and 11.11)**.**

(i) A dense set of coefficients. *Let $\alpha = \frac{1}{2}\sum_i \log_{m_i} q_i$ for finitely many (at least one) pairs of integers $m_i \geq 2$ and $q_i = 2^{r_i}$ with $r_i \geq 1$ and q_i dividing m_i; these numbers form a dense subset of $(0,\infty)$ that contains every positive rational and some transcendental numbers, such as $\frac{1}{2}\log_6 2$. For each such α there are a fixed Boolean constraint game and a fixed linear system game with $d^{\rightarrow}(\varepsilon) = \alpha\log(1/\varepsilon) + O(\log\log(1/\varepsilon))$ for small ε, minimum Schmidt rank of a pure strategy $\Theta(\varepsilon^{-\alpha})$, and minimum pure entropy and formation $\alpha\log(1/\varepsilon) + O(1)$. For rational α the games can be chosen so that, over*

pure strategies, the coefficient of $\log(1/\varepsilon)$ *in the minimum Rényi entropy* H_β *is* α *for* $0 \le \beta \le 1$ *and* $\min\{\alpha, \beta/(\beta-1)\}$ *for* $1 < \beta < \infty$ *(Theorem 11.7, which also treats* H_∞*, smooth min-entropy and one-shot concentration).*

(ii) Almost inverse loss. *The configuration-lamp game of [101] has* $d^{\to}(\varepsilon) = \varepsilon^{-1+o(1)}$*, with an explicit lower bound* $c\,\varepsilon^{-1}2^{-4(24+\log\log(1/\varepsilon))^2}$ *and upper bound* $O(\varepsilon^{-1}\log(1/\varepsilon))$*.*

(iii) The Slofstra–Vidick baseline. *Slofstra and Vidick bounded the least dimension of their game [134] between* $\varepsilon^{-1/6}$ *and* $\varepsilon^{-1/2}$*; our companion paper shows that it is* $\Theta(\varepsilon^{-1/2})$ *and that the source has profile* $\Theta(1/e)$ *at every visibility [101, Theorem F]. The game's least pure entropy and formation are* $\frac{1}{2}\log(1/\varepsilon) + O(1)$*.*

(iv) A polynomial window. *A fixed Boolean constraint game built from Pauli lamps on the points of Houghton's group* H_3 *has* $c\,\varepsilon^{-\beta} \le d^{\to}(\varepsilon) \le C\varepsilon^{-1/4}$*, where* $\beta = (2\log_2\lambda+1)^{-1} = 0.0734595360384268649\ldots$ *and* λ *is specified in Proposition 13.6a.*

By Theorem B each holds in both directions.

In (i) the source is a finitely presented group $L_{q,m}$ with a dilation $xyx^{-1} = y^m$ and q lamps cyclically permuted by y: the spectral law of y is nearly invariant under multiplication by m and rotation by $1/q$. Every invariant law satisfies a Frostman bound of exponent $s = \log_m q$, hence has dimension at least s. This fixes the minimum spectral scale relevant to the profile, and matching approximate models attain the corresponding rank and entropy rates. The exponents of independent copies add, without a joint spectral measure (Theorems 10.2 and 10.9). The usable-entanglement coefficient comes from a triangular extraction on the original state with cyclic labels (Theorem 6.2 and Lemma 6.3). The transcendence of $\frac{1}{2}\log_6 2$ is the Gelfond–Schneider theorem. In the other direction, the logarithmic coefficient of a fixed game of value one, when it exists, is left-c.e. relative to the second Turing jump $0''$ (Theorem 11.14). Every positive Δ^0_2 real occurs with a matched original-mixed-state usable law [57, Theorem 16.6 and Corollary 16.7]; the full usable spectrum remains open (Problem 18.3). The positive pure original-state entropy spectrum is exactly the positive Δ^0_2 reals (Theorem 11.16). In (ii) the source has 2^N Pauli lamps whose pairwise commutators have area $2^{O(\log^2 N)}$; the companion paper counts them in an arbitrary state, with errors summed in squares, and an extraction that counts Bell-error entropy rather than fidelity gives the usable rate (Theorems 3.1 and 6.4). In (iv) Houghton's far commutators have area $O(R^{\log_2\lambda})$ (Proposition 13.6a), by an explicit six-type refinement of the filling method of [55]; on the original state a filling is paid only linearly in its area and length (Theorem 13.5), and this gives the lower exponent, while an explicit finite model that spreads one collective defect over a torus of clocks gives the upper one (Proposition 13.9). Whether the demand in (iv) has a limiting exponent, and what it is, remain open. On a calibrated native class of strategies the point game has demand $\Theta(\varepsilon^{-1/3})$, a conditional class law (Theorem 13.11; see Section 1.6).

The least usable entanglement of an original state carrying M pairwise approximate Pauli pairs at error $0 \le \kappa \le 1/100$ is of order $\min\{M, \kappa^{-2}\}$, and exactly M at $\kappa = 0$ (Theorem 6.12). The restriction $\kappa \le 1/100$ cannot be dropped: at $\kappa = 2$ a product state satisfies every defect bound and has rate zero (end of Section 6.7). The spectrum refinements in Section 11.6 separate computable pure-state optimization, convex-roof formation and physical one-way leakage; Appendix F gives quantitative obstructions and sufficient conditions for the logarithmic floor. For the collective point model, the magnetic-cycle estimate fixes the best defect order obtainable by subdividing its fixed holonomy (Proposition 13.9a), without settling the quarter lower bound.

What a game sees through approximate relations, an exact source sees through exact ones. Under exact relations the demand is governed by a single growth function. These results

belong to the exact-source contract (EX) of Section 2.8, not to finite-game loss (Section 1.6).

Theorem D (The price of an address; Theorems 14.1, 14.3, 14.5 and 14.7 to 14.9)**.** *If Bob satisfies the relations of a source exactly, in the sector $J = -I$ of a necessarily infinite-dimensional representation, and Alice mirrors each generator with error e on an arbitrary original mixed state, then the demand in both directions is at least a fixed positive multiple of the number of independent Pauli qubits addressable by words of length c/e. For Pauli lamps on the points of H_3, on the k-subsets of its points, and on its configurations, the demand is $\Theta(1/e)$, $\Theta_k(e^{-k})$ and $2^{\Theta(1/e)}$, with matching finite-rank constructions; a five-generator replication source is also exponential. With a central reflection of visibility Δ instead of $J = -I$, the subset sources have, uniformly for $e \le c_k\Delta$, the joint law $\Theta_k(\Delta^{k+2}e^{-k})$.*

Exactness can be exponentially expensive. The approximate configuration models of [101, Proposition 9.4] cost only $O(e^{-2}\log(1/e))$ ebits, the exact ones $2^{\Theta(1/e)}$, so a correction that keeps the original state and its resource needs error $\Omega(1/\log(1/e))$; the collective point models cost $O(e^{-1/2})$ ebits against $\Theta(1/e)$ for exact ones (Proposition 15.7).

For near-optimal Bell violations, finite-dimensional nonattainment alone need not force divergent usable entanglement. The I_{3322} Bell functional is not attained in finite dimension, as Pauwels [119, Theorem 2], Coladangelo [37, Theorem 11.1] and the first author [53, Theorem 1] showed independently; it is attained by an infinite-dimensional spatial strategy [37, Theorem 11.1], [53, Theorem 2], and its least dimension at Bell gap ε grows like $\log(1/\varepsilon)$ [53, Theorem 6]. The Schmidt coefficients of the attainer of [53] have exponential tails [53, Section 5.2], and bounded usable entanglement is an elementary consequence of those tails. The analytic proofs of attainment and of the tails are not formalized, so we state the consequence conditionally.

Theorem E (A conditional Bell-gap contrast; Theorem 16.1)**.** *Conditional on the spatial attainment theorem and the tail estimate of [53], strategies for the I_{3322} functional with Bell gap tending to zero need only bounded usable entanglement, in either direction.*

This changes both the contract and the class (Section 1.6), so it does not decide whether Theorem A extends beyond classes (S) and (C). What the contrast does show is a difference in the kind of nonattainment. The infinite-dimensional construction of the I_{3322} optimum and the optimality conjecture are due to Pál and Vértesi [116]; the cited 2026 papers concern its analytic resolution. In the classes of Theorem A a normal perfect strategy of any entropy already has a finite perfect block, by a theorem of Kim, Paulsen and Schafhauser [83] and its standard analogue for constraint games (Remark 16.2), whereas, by [53, Theorem 2 and Section 5.2], the I_{3322} optimum is attained by a normal strategy of finite entropy.

The engine. Behind these results is a set of tools that keep the state; the first two are proved in our companion paper [101] and are used here as stated there. In the hyperlinear-profile route of Slofstra and Vidick, the passage from a general strategy to a normalized-trace model, through a spectral projection of the reduced state, changes the precision from order $\sqrt{\varepsilon}$ to order $\varepsilon^{1/4}$ [134, Theorem 5.1 and Lemmas 5.5, 5.6], and rounding to a maximally entangled strategy [112] replaces the state. Our arguments keep the state-weighted model through these steps. Original-state self-testing and mixed-state assumption lifting were discussed above; what differs here is the nonattaining fixed-game regime and its resource laws. (For the memory of a repeatedly used channel, the weighted one-round theorem of our earlier paper already works with a non-flat state [54, Theorem 10.1]; the tools below do this for both players of a game.)

- *A qubit count in an arbitrary state* (Theorem 3.1, from [101, Theorem C]): pairwise relations among $2M$ approximate Pauli involutions, measured in the state and summed in squares,

together with their commutators with the square root of the state, bound its entropy below by about M.

- *Transfer and return with the state kept* (Theorems 4.3 and 4.5, from [101, Theorem D], and Theorem 4.9): the cost of a linear or Boolean constraint game is a state-weighted profile of its algebra at precision $\Theta(\sqrt{\varepsilon})$, in both directions, with the whole spectrum of the state preserved.
- *A retained-state compiler* (Theorem 5.6): every finite presentation with a central involution is compiled into a Boolean constraint game of arity three, with a split algebraic retraction that keeps the whole state; corner maps (Proposition 5.5) recover a group inside a game algebra without cutting out a sector.
- *Extraction on the original mixed state* (Section 6): triangular and signed regular extraction, aggregate and finite-group versions with a center-aware one-way protocol, and Alice-local recovery for the forward direction.
- *A linear filling charge* (Theorem 13.5): on the original state, a null word of length n with a filling of A cells has error $O((A+n)e)$ at precision e, although the conjugators of its cells can be long.

The tools added here are the Boolean case of the correspondence, the compilers and the original-state extraction; they turn the entropy and dimension bounds of the companion into bounds on usable entanglement.

1.4 The ideas of the proofs

Precision lives in the state. If Bob's reduced state is ρ and $\xi = \rho^{1/2}$, a strategy of loss ε gives Bob operators that satisfy the relations of the game up to $\sqrt{\varepsilon}$ in the norm $\|Y\xi\|_2 = \mathrm{Tr}(Y^*Y\rho)^{1/2}$, and that nearly commute with ξ, again to order $\sqrt{\varepsilon}$: consistency between Alice and Bob forces it (Lemma 2.2). Fixing a state is not like fixing a trace: conjugation by a unitary moves the state, and the price is the commutator with ξ. Every estimate in this paper is made in such a norm, and the only cost of changing the state is a commutator of this kind. Errors are summed in squares wherever possible, because the counting and extraction arguments below charge each qubit only the sum of squares of its pair errors.

Counting qubits without flattening. Exact Pauli pairs X_i, Z_i force dimension 2^M. To see approximately the same in a state, the companion paper's argument [101, Section 4] proceeds as follows: embed the space coherently into a register indexed by $\{0,1\}^M$ through the products $Z_1^{g_1}\cdots Z_M^{g_M}$; on the register the shifts and signs are exact Pauli operators, and the embedding intertwines them with the X_i, Z_i up to an error that the two-sided twirl of Lemma 2.9 computes as a sum of squares of pair relations; pinching the embedded state then gives about one bit per pair (Theorem 3.1).

From a count to Bell pairs. A count bounds the entropy of Bob's marginal, which a separable state can also have. To get usable entanglement, both players apply the coherent embedding, each to their own operators; actual mirrors then make the two exact label registers nearly agree, and every Bell stabilizer of the output fails with probability $O(M\kappa^2)$ (Theorem 6.4). The decisive point is to use these failures through their *entropy*: after a Pauli twirl the output is Bell-diagonal, and hashing gives $M - \sum_k h(p_k)$ ebits however the errors are correlated. A fidelity bound for the whole register would add the failures and lose a factor M.

Why unreachable perfection is enough. In Theorem A there is no rate and no explicit source. The tracial limit of Bob's moments is a finite von Neumann algebra. When the game has no finite-dimensional perfect strategy, the limit has no type I piece, since such a piece would yield one; so the limit is of type II and contains every matrix algebra. Pauli matrices there are approximated, in the trace norm, by bounded polynomials in the game's measurements; evaluated on the original strategies, on both sides, and rounded locally, they are approximate Pauli pairs with small mirrors, and the fixed-size extraction lemma (Lemma 6.5) applies.

Why the directions agree. Alice can map her half of a purification to a conjugate copy of Bob's side; on the output the two parties hold mirror images, so the two one-way rates coincide, and they are no larger than before. To keep playing an agreement game, Alice measures the image of her old measurements under the correlation map S; the new success is $\langle w, H^3 w\rangle$ for the self-adjoint contraction H whose quadratic form is the old success, and $h^3 \geq 3h - 2$ on $[-1, 1]$ gives the factor three.

Rates from transport. For a source with $xyx^{-1} = y^m$ and lamps rotated by y, the relations make the spectral law of y, measured in the state, nearly invariant under multiplication by m and under rotation by $1/q$. A nearly invariant law is close to an invariant one, and every invariant law satisfies a Frostman bound of exponent $\log_m q$, hence has dimension at least $\log_m q$; so a law with few atoms is far from all of them. Entropy, rather than dimension, needs the commutator of y with ξ once more: it lets both halves of a purification read the same spread spectral law (Theorem 9.6). For several copies we cannot use a joint spectral measure; instead we measure spectral arcs of short words in reverse order of depth, so that the errors form geometric series (Theorem 10.9).

Rates from areas. For lamps on configurations or points of Houghton's group, the qubits come from transported Pauli pairs, and the cost of making two transported pairs commute is governed by the area of a filling. On a non-flat state each cell of a filling is reached through a conjugator, and charging every cell its own conjugator costs area times length. Cutting the van Kampen diagram into a tree of faces, and substituting the whole boundary of each subdiagram at once, shares the conjugators, so a filling costs only linearly in its area and length (Theorem 13.5). Quasi-polynomial areas give the configuration-lamp rate, and polynomial areas the lower end of the point window.

Exact sources and collective models. When the relations are exact, every independent qubit is free, and the cost is the number of qubits addressable by short words. Approximate models are cheaper because they can merge addresses or, as in the collective model of the point game, replace a sum of many small defect logarithms by the logarithm of their product, which is a single bounded projection. The difference can be exponential (Proposition 15.7).

1.5 How the results fit

Table 1 lists the main results with the resource each controls and the mechanism behind it. Read down the table, the picture is as follows. Theorem A is qualitative and uses only the absence of finite-dimensional representations of the game algebra: type II limits contain arbitrarily large matrix algebras, and the actual mirrors pull them back to the original state. The rates need quantitative invariants of a source, and each invariant enters through one tool. The Frostman exponent of the measures invariant under the dilation and the rotation gives the logarithmic coefficients, a dense set containing every positive rational; areas of

commutators in Houghton's group give the configuration-lamp rate and the point window; and the growth of short addresses gives the exact-source laws. The conversions of errors behind every rate, from loss to mirrors, relation errors, Bell-stabilizer failures and ebits, are collected once in Section 2.8. For agreement games, which include class (C), each directional demand at loss ε is at least the other at loss 3ε (Theorem 8.2); for general games in class (S), Proposition 8.4 gives only the weaker square-root repair, with $\varepsilon + C_G\sqrt{\varepsilon}$ in place of 3ε.

object	resource	law	mechanism
games in classes (S), (C)	usable E_{D}, both directions	divergent iff no finite perfect strategy	type II limit, bounded lifts, actual mirrors
games with a sign and an even dilation (e.g. Abels' group)	E_{D}	at least $\frac{\log(1/\varepsilon)}{4\log m} - O(\log\log\frac{1}{\varepsilon})$	actual-mirror decoder, triangular extraction
agreement games	$E_{\mathrm{D}}^{A\to B}$ vs $E_{\mathrm{D}}^{B\to A}$	equal up to loss factor 3, optimal for Bob-fixed local symmetrization	reflected state, KMS correlation maps
lamp-sum families (Boolean constraint and linear games), α in a dense set	E_{D}; pure Schmidt rank; Rényi (α rational)	$\alpha\log\frac{1}{\varepsilon} - O(\log\log\frac{1}{\varepsilon})$; $\Theta(\varepsilon^{-\alpha})$; finite-order spectrum	dilation $\times m$ and rotation $\frac{1}{q}$: Frostman exponent $\log_m q$
configuration-lamp game	E_{D}	$\varepsilon^{-1+o(1)}$	2^N lamps, quasi-polynomial pair areas
Slofstra–Vidick game	dimension; entropy	$\Theta(\varepsilon^{-1/2})$; $\frac{1}{2}\log\frac{1}{\varepsilon} + O(1)$	doubling and half-turn: Haar law
point-Pauli game	E_{D}	between $\varepsilon^{-\beta}$ and $\varepsilon^{-1/4}$, $\beta = \frac{1}{2\log_2\lambda+1}$	polynomial far-commutator area, linear filling charge; collective model
exact sources	E_{D}, both directions	$\Theta(\frac{1}{e})$, $\Theta(e^{-k})$, $2^{\Theta(1/e)}$	growth of short addresses
I_{3322} (Bell gap)	E_{D}	bounded near the Bell optimum (conditional on [53])	normal attainer with geometric tails

Table 1: The main results. ε is a game loss from value one, except in the I_{3322} row, where it is a gap below the Bell optimum; e is a source precision; all games are fixed before ε varies.

1.6 Scope and status

We state the scope once here, and the later sections refer to this subsection. Three distinctions run through the paper. *Original state versus replacement state*: every lower bound on usable entanglement is a bound on the distillable entanglement of the original mixed state, with its purifying environment untouched; bounds on marginal entropy, pure Schmidt rank, entanglement of formation or dimension are stated as such and are never used as bounds on E_{D}. *Loss versus Bell gap*: ε is a game loss from value one, except in Section 16, where a gap below the Bell optimum of I_{3322} is used, which no affine rescaling turns into a game loss (Section 16.1). *Unrestricted versus class laws*: a demand is an infimum over all finite-dimensional strategies; the native laws of Section 13.5 are laws on a restricted, calibrated class of strategies, and the exact-source laws of Section 14 are laws for a different contract, exact source relations and approximate mirrors.

The companion paper [101] proves the qubit count, transfer and return for linear system games, the CHSH(n) and Slofstra–Vidick rates, and the explicit configuration-lamp game

with its entropy, formation and dimension bounds. We state the inputs we use with precise pointers in Appendix D. The realization and matched coefficient laws used in Section 11.6 and Problem 18.3 are supplied by [57, Theorem 1.1, Theorem 16.6 and Corollary 16.7]. Proofs identify the imported results at their uses; the remaining arguments are written in the text or appendices. The native class laws of Theorems 13.11 and 13.12 are *conditional* on estimates from our manuscript on immediate delivery [56] and on an aggregate estimate, which have not been proved self-containedly. The specifications of the restricted linear compiler for several copies (Theorem 11.3: arity, amplification, algorithm and exports) are *conditional* on a self-contained proof of that compiler; the existence of linear system games with every positive rational exponent is not (Corollary 11.6), and neither are the Boolean constraint family and the one-copy linear family. The I_{3322} statement Theorem 16.1 is *conditional* on the spatial attainment theorem and tail estimate of [53]. Table 2 identifies the proof locations and computational scope of the results; the section How this work was done describes the process and the accompanying finite checks.

Slofstra and Vidick's proof of [134, Proposition 6.2] takes from [132, Proposition 4.8] a $2d$-dimensional lift in which the marked involution is balanced. With two or more non-involutive generators, as in the group behind the Slofstra–Vidick game, the literal lift is unbalanced, so that step does not apply as written. This affects only the dimension factor of the upper bound: their value one, their lower bounds and the non-triviality of J, which they take from [132, Lemma 5.2], do not use the lift. Our linear games therefore use a balanced construction (Proposition 5.10), and our companion paper proves the lift for the Slofstra–Vidick group [101, Lemma 11.13], so [134, Theorem 1.1] holds as stated (Section 9.4).

The results are compositions of established tools with new accounting, and we say so at each result: tracial and synchronous correspondences [34, 118, 83], Kaplansky density, Pauli and regular-representation extraction [66, 107], Petz and pretty-good recovery [120, 9, 17, 73, 123], the uncertainty relation with memory [16] and hashing [51, 14]. Bounded literature searches found no exact preemption of the composed headline statements; this is not an exhaustive priority verdict, and we make no claim of being first.

We do not claim a classification of all games or an exact description of the full mixed-state usable-entanglement coefficient spectrum. The positive pure original-state spectrum is classified by Theorem 11.16. We do not claim a device-independent protocol, efficient decoders, a matching quarter exponent for the point game, or a logarithmic floor for every nonattaining game; these and other open targets are stated in Section 18. Constants are deliberately poor and depend on the fixed game.

1.7 Related work and credits

The Hilbert–Schmidt picture of a shared state, with the reduced state as a factor ξ, is the standard form of the state; Slofstra and Vidick use it for general strategies [134], and Marrakchi and de la Salle use it to round near-synchronous strategies [97]. The coupling of two spectral measures in one state is Connes' joint distribution [41]. Defect bounds in the state of a near-perfect strategy of a boolean constraint system game, and embeddings of group algebras into such games with finite certificates, are due to Fanizza, Kroell, Mehta, Paddock, Rochette, Slofstra and Zhao [60], whose relation-decomposition bound is Lemma 3.9 of Mehta, Slofstra and Zhao [99]; Slofstra's finite gadgets and embedding theorem [132, 133] and Paddock's constraint-system machinery [112] underlie our compiler, whose new part is the split retraction. The finite-dimensional structure of perfect synchronous strategies is due to Paulsen, Severini, Stahlke, Todorov and Winter and to Kim, Paulsen and Schafhauser [118, 83]; approximate tracial rounding to Vidick, Marrakchi and de la Salle, Paddock, and Culf [145, 97, 112, 45]; Pauli stability for synchronous strategies to de la Salle [49]. Thoma's classification of the characters of the infinite symmetric group [136], Cornulier's criterion for finitely presented

wreath products [48], and Johnson's presentation of Houghton's group [77, 84] are used for the exact sources. The functional-inequality framework behind the Rényi spectrum is that of [7, 126], and one-shot concentration is due to [144, 90, 25, 109]. The undecidable core of Theorem 7.6 is Slofstra's [132]; the remark after that theorem credits the related game disjunction of Mančinska and Schmidt [95] and compares the theorem with the RE-completeness of entangled constraint satisfaction problems [46].

The configuration lamps come from our arXiv paper [54]. There, copies of S_3 sit on the configurations of the points of Houghton's group; the paper bounds the sofic profile of the resulting group and the dimension of its unitary models at normalized-trace precision e, which is at least $\exp(c\,e^{-2}(\log(1/e))^{-2\kappa})$ with $\kappa = 6 + \log 107$ at fixed visibility [54, Theorem A], and its weighted one-round theorem bounds the entropy of a non-flat memory state through transported copies of S_3 [54, Theorem 10.1]. The companion paper [101] puts Pauli pairs on the same configurations, builds the game and bounds the entropy, formation and dimension of its near-winners. Neither paper bounds the distillable entanglement of a strategy for a game; that is what Theorem 12.4 adds.

1.8 Organization

For the qualitative law, read Section 2.8, the fixed-size extraction lemma (Lemma 6.5) and Section 7. The logarithmic laws need in addition the compiler and extraction tools of Sections 5 and 6, the lamp sources of Section 10 and the coefficient family of Section 11. The polynomial demands are in Sections 12 and 13. Section 14 changes to the exact-source contract (EX) of Section 2.8.

Section 2 fixes norms, the standard form, a calculus for words in a state, and the resources and contracts. Part I, the engine, consists of Section 3 (the qubit count of the companion), Section 4 (transfer and return for linear and Boolean constraint games), Section 5 (compilers) and Section 6 (extraction). Part II, the law, is Section 7 (Theorem A and the results around it) and Section 8 (directions). Part III, the spectrum, consists of Section 9 (the Slofstra–Vidick baseline), Section 10 (lamp sources), Section 11 (a dense set of coefficients and the arithmetic of coefficients), Section 12 (the configuration-lamp game), Section 13 (the point game) and Section 14 (exact sources). Part IV contains the obstructions (Section 15), the conditional Bell-gap contrast (Section 16), the logarithmic floor on a class and a sufficient criterion for a counterexample (Section 17) and open problems (Section 18), followed by a discussion and an account of how the work was done. The appendices give the proof of the white-box protocol, a table of the status of every result, reproducibility commands and computational scope, the results used from our companion papers, a tapered reservoir for the game of Ji, Leung and Vidick, the candidate classes and method obstructions for the logarithmic floor, and, for the point window, the six-type far-commutator fillings, the fixed-holonomy subdivision estimate and an obstruction to raw cross-copy estimates.

2 Preliminaries

Unless stated otherwise, strategies are finite-dimensional. The exceptions are stated where they occur: the Alice-local recovery and the local symmetric output accept normal states (Sections 6.5 and 8.1), Section 16 treats normal spatial strategies, and the exact source representations of Section 14 have infinite-dimensional ambient spaces. This section fixes the norms in which every error is measured, the dictionary between a shared pure state and Hilbert–Schmidt operators on one side, and a short calculus for words in unitaries measured in a fixed state. The last three subsections recall the projection-product inequality, which makes twirls of Hilbert–Schmidt vectors useful, the passage from a state to its restriction

to a sector, and the facts about relative entropy that the counting argument needs. Except for the polar correction (Lemma 2.7) and the transfer across the parties (Lemma 2.11), the lemmas of this section are proved in our companion paper [101, Section 3], in the same or a stronger form; we state them here, with a pointer, and do not repeat their proofs.

2.1 Norms, states and resources

For an operator T between Hilbert spaces, $\|T\|_2 = (\operatorname{Tr} T^*T)^{1/2}$ is the unnormalized Hilbert–Schmidt norm, $\|T\|_1$ the trace norm and $\|T\|_{\mathrm{op}}$ the operator norm. We use $\|XTY\|_2 \le \|X\|_{\mathrm{op}}\|T\|_2\|Y\|_{\mathrm{op}}$, Hölder's inequality $\|ST\|_1 \le \|S\|_2\|T\|_2$, and the invariance of $\|\cdot\|_2$ under multiplication by unitaries on either side. On $\mathbb{C}^d$, the normalized norm is $\|T\|_f = \|T\|_2/\sqrt{d}$. An *involution* is a Hermitian unitary. We write $[S,T] = ST - TS$ and $\{S,T\} = ST + TS$ for operators; for words, $[u,v] = uvu^{-1}v^{-1}$ is the group commutator.

Let ρ be a density matrix on $\mathbb{C}^d$ and $\xi = \rho^{1/2}$, so that $\|\xi\|_2 = 1$. For an operator Y on $\mathbb{C}^d$ we call

$$\|Y\xi\|_2 = \operatorname{Tr}(Y^*Y\rho)^{1/2} \qquad \text{and} \qquad \|[Y,\xi]\|_2$$

the *state-weighted norm* of Y and the *state commutator* of Y. For the flat state $\rho = I/d$ the first is $\|Y\|_f$ and the second vanishes. We use without comment that $\|XY\xi\|_2 \le \|X\|_{\mathrm{op}}\|Y\xi\|_2$, that $\|Y\xi U\|_2 = \|Y\xi\|_2$ for unitary U, that $\|\xi Y\|_2 = \|Y^*\xi\|_2$, that $\|(U^*-I)\xi\|_2 = \|(U-I)\xi\|_2$ for unitary U, and that $[U^*,\xi] = -U^*[U,\xi]U^*$, so $\|[U^*,\xi]\|_2 = \|[U,\xi]\|_2$. For contractions S,T the Leibniz rule $[ST,\xi] = S[T,\xi] + [S,\xi]T$ gives $\|[ST,\xi]\|_2 \le \|[S,\xi]\|_2 + \|[T,\xi]\|_2$.

Logarithms are to base two unless written ln, $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy, S is the von Neumann entropy in bits, and $D(A\|B) = \operatorname{Tr} A(\log A - \log B)$ is the relative entropy of positive operators A, B with $\operatorname{supp} A \subseteq \operatorname{supp} B$.

Resources. A unit vector $\psi \in H_A \otimes H_B$ has a Schmidt decomposition $\psi = \sum_{k=1}^{r} \sqrt{\lambda_k}\, e_k \otimes f_k$ with $\lambda_k > 0$ and orthonormal families (e_k), (f_k). Its reduced density matrices ρ_A and ρ_B have the same nonzero eigenvalues λ_k. We keep four quantities apart: the *Schmidt rank* $r = \operatorname{rank}\rho_B$, which is at most the local dimensions; its logarithm; the *entanglement entropy* $S(\rho_A) = S(\rho_B) \le \log r$, the number of ebits of a pure state; and, for a mixed state σ on $H_A \otimes H_B$, the *entanglement of formation* [14]

$$E_{\mathrm{F}}(\sigma) = \inf\Big\{\sum_j p_j S(\operatorname{Tr}_B |\psi_j\rangle\langle\psi_j|) \,:\, \sigma = \sum_j p_j |\psi_j\rangle\langle\psi_j|\Big\}.$$

The distillable entanglement $E_{\mathrm{D}}(\sigma)$ is at most $E_{\mathrm{F}}(\sigma)$, so a lower bound on E_{F} says nothing about E_{D}.

For a nonlocal game G with quantum value $\omega_q(G)$, $\mathcal{E}(G,\varepsilon)$ is the least d such that some strategy with $H_A = H_B = \mathbb{C}^d$ wins with probability at least $\omega_q(G) - \varepsilon$ [134]. Compressing each player's measurements to the support of the reduced state does not change any probability [101, Remark 2.9], so $\mathcal{E}(G,\varepsilon)$ is also the least Schmidt rank of such a strategy. A lower bound on the Schmidt rank of every strategy that wins with probability at least $\omega_q(G) - \varepsilon$ is therefore a lower bound on $\mathcal{E}(G,\varepsilon)$; a lower bound on the entanglement entropy is a lower bound on $\log \mathcal{E}(G,\varepsilon)$.

2.2 The standard form

A shared pure state turns operators of either player into Hilbert–Schmidt operators on one side, with the reduced state appearing as a factor ξ on the left or on the right.

Lemma 2.1 (Standard form)**.** *Let $\psi \in H_A \otimes H_B$ be a unit vector, $\rho = \rho_B$ and $\xi = \rho^{1/2}$.*

(a) *For every operator Y on H_B, $\|(I\otimes Y)\psi\| = \|Y\xi\|_2$.*

(b) *For every Hermitian operator A on H_A there is a Hermitian operator $\tilde{A}$ on H_B with $\|\tilde{A}\|_{\mathrm{op}} \le \|A\|_{\mathrm{op}}$ such that, for every operator Y on H_B,*

$$\|Y\xi - \xi\tilde{A}\|_2 \;\le\; \|(A\otimes I - I\otimes Y)\psi\|.$$

The same holds with the roles of the two factors exchanged.

This is [101, Lemma 3.1]. It is the standard form of the state: Slofstra and Vidick start their analysis of general strategies from the identity $\langle\psi,(B^T\otimes C)\psi\rangle = \mathrm{Tr}(B\lambda C\lambda)$, $\lambda=\rho^{1/2}$ [134, proof of Proposition 5.4], and Marrakchi and de la Salle use the standard form in the same way for rounding near-synchronous strategies [97].

2.3 Consistency gives state commutation

If an observable of one player agrees with a Hermitian operator of the other on the shared state, the observable nearly commutes with the square root of the reduced state, with no loss in the exponent.

Lemma 2.2. *Let B be an involution and T a Hermitian operator on $\mathbb{C}^d$, of any norm, with $\|B\xi - \xi T\|_2 \le \theta$. Then $\|[B,\xi]\|_2 \le \sqrt{2}\,\theta$.*

This is [101, Lemma 3.2], the Cauchy–Schwarz argument of [134, (5.1)] with an explicit constant and with a Hermitian partner of arbitrary norm; the constant $\sqrt{2}$ cannot be improved.

2.4 State-weighted calculus

Let S be a finite alphabet and $\varphi : F(S) \to \mathrm{U}(d)$ a homomorphism from the free group, so that $\varphi(w)$ depends only on the free reduction of a word w. The defect of a relator in the state is $\|(\varphi(r)-I)\xi\|_2$. Unlike the normalized trace, a fixed state does not ignore conjugation, so conjugators are charged through state commutators.

Lemma 2.3. *Suppose $\|[\varphi(s),\xi]\|_2 \le e_c$ for every $s\in S$.*

(a) *$\|[\varphi(w),\xi]\|_2 \le |w|\,e_c$ for every word w.*

(b) *For unitaries $U_1,\dots,U_k$, $\|(U_1\cdots U_k - I)\xi\|_2 \le \sum_l \|(U_l - I)\xi\|_2$.*

(c) *For unitaries g, U, $\|(gUg^* - I)\xi\|_2 \le \|(U-I)\xi\|_2 + 2\|[g,\xi]\|_2$.*

(d) *If w is freely equal to $\prod_{l=1}^{k} u_l r_l^{\pm 1} u_l^{-1}$ for words u_l, r_l with $\|(\varphi(r_l)-I)\xi\|_2 \le e$ for all l, then $\|(\varphi(w)-I)\xi\|_2 \le \sum_{l=1}^{k}(e + 2|u_l|e_c)$.*

(e) *For contractions $Y_1,\dots,Y_m$ and unitaries $Y_1',\dots,Y_m'$,*

$$\|(Y_1\cdots Y_m - Y_1'\cdots Y_m')\xi\|_2 \;\le\; \sum_{l=1}^{m}\Big(\|(Y_l - Y_l')\xi\|_2 + 2\sum_{l'>l}\|[Y_{l'}',\xi]\|_2\Big).$$

This is [101, Lemma 3.3]; parts (b), (c) and (e) do not use the hypothesis.

Part (d) needs short conjugators. For relations with bounded area the next lemma supplies them; it is the usual relation between area and conjugator length, read off a van Kampen diagram [92, Chapter V, Section 1], [22, Section 4]. Let R be a finite set of words. A word w has *area at most k in R* if it is freely equal to a product of k conjugates of elements of $R^{\pm 1}$.

Lemma 2.4 (Conjugator lengths)**.** *Let every word of R have length at most K. If w has area at most k in R, then w is freely equal to a product $\prod_{l=1}^{k'} u_l r_l^{\pm 1} u_l^{-1}$ with $k' \le k$, $r_l \in R$ and $|u_l| \le |w| + K(k+1)$ for all l.*

Our companion paper proves the sharper bound $|u_l| \le |w| + kK$ [101, Lemma 3.4].

Lemma 2.5 (Rounding). *For a unitary U let $\mathrm{r}(U)$ be the involution obtained by replacing each eigenvalue λ of U by 1 if $\operatorname{Re}\lambda \ge 0$ and by -1 otherwise. Then*

$$\|(\mathrm{r}(U) - U)\xi\|_2 \le \|(U^2 - I)\xi\|_2 \qquad \text{and} \qquad \|[\mathrm{r}(U), \xi]\|_2 \le \|[U, \xi]\|_2 + 2\|(U^2 - I)\xi\|_2.$$

This is [101, Lemma 3.6].

Lemma 2.6 (Rotating the state). *For a unitary U, $\|U^*\rho U - \rho\|_1 \le 2\|[U, \xi]\|_2$.*

This is [101, Lemma 3.7].

Operators that are unitary only approximately, in the state, are corrected by their polar part; both one-sided defects are needed.

Lemma 2.7 (Polar correction). *Let A be an operator on $\mathbb{C}^d$ and $\xi \ge 0$. Every unitary U with $A = U|A|$ satisfies*

$$\|(U - A)\xi\|_2 \le \|(A^*A - I)\xi\|_2, \qquad \|\xi(U - A)\|_2 \le \|(AA^* - I)\xi\|_2,$$

and hence $\|[U, \xi]\|_2 \le \|[A, \xi]\|_2 + \|(A^*A - I)\xi\|_2 + \|(AA^* - I)\xi\|_2$.

Proof. A square matrix has a polar decomposition $A = U|A|$ with U unitary; let U be any such unitary. For $t \ge 0$, $|1 - t| \le |1 - t^2|$, so $(I - |A|)^2 \le (I - A^*A)^2$ and $\|(U - A)\xi\|_2^2 = \|(I - |A|)\xi\|_2^2 = \operatorname{Tr}(\xi(I - |A|)^2\xi) \le \|(A^*A - I)\xi\|_2^2$. Next $A^* = |A|U^* = U^*|A^*|$, since $|A^*| = (U|A|^2U^*)^{1/2} = U|A|U^*$; this is a polar decomposition of A^* with the unitary U^*, so the first bound gives $\|(U^* - A^*)\xi\|_2 \le \|(AA^* - I)\xi\|_2$, and $\|\xi(U - A)\|_2 = \|(U^* - A^*)\xi\|_2$. The last bound follows from $[U, \xi] = [A, \xi] + (U - A)\xi - \xi(U - A)$. □

For the flat state this is the standard polar rounding [134, Lemma 2.8(a)]. The left defect cannot be dropped: for $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ and $\xi = \operatorname{diag}(1, 0)$ we have $(A^*A - I)\xi = 0$, but every unitary U with $A = U|A|$ has $\|\xi(U - A)\|_2 = 1$.

2.5 Hilbert–Schmidt twirls

Lemma 2.8 (Projection products; [55, Lemma 4.1], [101, Lemma 3.8]). *Let $P_1, \dots, P_k$ be orthogonal projections on a Hilbert space. Every vector y satisfies*

$$\|y\|^2 - \operatorname{Re}\langle y, P_k \cdots P_1 y\rangle \;\le\; 2\sum_{l=1}^{k} \|(I - P_l)y\|^2.$$

The quantum union bounds for sequential projective measurements control instead the probability that all outcomes occur, for example $1 - \|P_k \cdots P_1 y\|^2 \le 4\sum_l \|(I - P_l)y\|^2$ for a unit vector y [64, Theorem 1], with a version with a tunable parameter by Khabbazi Oskouei, Mancini and Wilde [82]; the twirls below use the real part of the overlap that the lemma bounds.

For an involution Z on $\mathbb{C}^d$, the map $\alpha_Z(v) = ZvZ$ on $M_d(\mathbb{C})$ with the Hilbert–Schmidt inner product is unitary and equal to its inverse, hence self-adjoint, *whatever state is in use.* So $A_Z^{\pm} = \frac{1}{2}(I \pm \alpha_Z)$ are orthogonal projections. The state enters only through the vectors that are twirled.

Lemma 2.9 (Two-sided twirls). *Let $Z_1, \dots, Z_k$ be involutions on $\mathbb{C}^d$, $v \in M_d(\mathbb{C})$, and fix signs $\sigma_j \in \{\pm 1\}$. For $g \in \mathbb{Z}_2^k$ put $s(g) = \prod_j \sigma_j^{g_j}$ and $v_g = Z_k^{g_k} \cdots Z_1^{g_1} v Z_1^{g_1} \cdots Z_k^{g_k}$. Then, for uniformly random g,*

$$\mathbb{E}_g \|s(g)v_g - v\|_2^2 \;\le \sum_{j:\sigma_j = +1} \|[Z_j, v]\|_2^2 + \sum_{j:\sigma_j = -1} \|\{Z_j, v\}\|_2^2.$$

This is [101, Lemma 3.9].

With a non-flat state, the map $T \mapsto \frac{1}{2}(T + ZTZ)$ is not an orthogonal projection for the weighted inner product $\mathrm{Tr}(\rho S^* T)$; twirling the vectors $v = Z_i\xi$, $X_i\xi$ and ξ avoids this.

2.6 Changing the state to a sector

The counting arguments compress the state, not the operators, to the eigenspace where a central involution acts as -1.

Lemma 2.10 (Sector states)**.** *Let J' be an involution on $\mathbb{C}^d$, $P_\pm = \frac{1}{2}(I \pm J')$, $p = \mathrm{Tr}(P_-\rho) > 0$, $\rho_- = P_-\rho P_-/p$, $\xi_- = \rho_-^{1/2}$ and $\eta = P_-\xi P_-/\sqrt{p}$.*

(a) $\|\xi_- - \eta\|_2 \le \|P_+\xi P_-\|_2/\sqrt{p}$, *and* $\|P_+\xi P_-\|_2 \le \frac{1}{2}\|[J', \xi]\|_2$.

(b) For every operator R, $\|R\xi_-\|_2 \le \big(\|RP_-\xi\|_2 + \|R\|_{\mathrm{op}}\|P_+\xi P_-\|_2\big)/\sqrt{p}$, *and* $\|R\xi_-\|_2 \le \big(\|R\xi\|_2 + \|R\|_{\mathrm{op}}\|[J', \xi]\|_2\big)/\sqrt{p}$.

(c) For every Hermitian contraction Y,

$$\|[Y, \xi_-]\|_2 \le \big(2\|[Y, P_-]\xi\|_2 + \|[Y, \xi]\|_2 + 2\|P_+\xi P_-\|_2\big)/\sqrt{p}.$$

For every unitary Y,

$$\|[Y, \xi_-]\|_2 \le \big(\|[Y, J']\xi\|_2 + 3\|[Y, \xi]\|_2 + \|[J', \xi]\|_2\big)/\sqrt{p}.$$

This is [101, Lemma 3.10].

For a Hermitian Y of arbitrary norm the proof of (c) gives the same bound with $2\|Y\|_{\mathrm{op}}\|P_+\xi P_-\|_2$ as the last term; we use (c) only for involutions and unitaries.

2.7 Entropy

We use three standard facts [101, Section 3.7]. If $\mathcal{P}(\sigma) = \sum_k P_k\sigma P_k$ is the pinching by orthogonal projections P_k with $\sum_k P_k = I$, then $\log \mathcal{P}(\sigma)$ is block diagonal, so $\mathrm{Tr}\,\sigma \log \mathcal{P}(\sigma) = \mathrm{Tr}\,\mathcal{P}(\sigma) \log \mathcal{P}(\sigma)$ and

$$S(\mathcal{P}\sigma) - S(\sigma) = D(\sigma\|\mathcal{P}\sigma) \ \ge 0. \tag{2.1}$$

The relative entropy does not increase under quantum channels [89, 140]. And $D(A\|B)$ is jointly convex in pairs of positive operators [87, 88]; since it is also positively homogeneous, $D(A_1 + A_2\|B_1 + B_2) \le D(A_1\|B_1) + D(A_2\|B_2)$.

2.8 Usable entanglement, precision and the two source contracts

One-way distillable entanglement. For a bipartite state ρ_{AB}, $E_{\mathrm{D}}^{A\to B}(\rho_{AB})$ is the largest rate, in ebits per input copy, at which local operations and classical communication from Alice to Bob alone convert independent copies of ρ_{AB} into maximally entangled pairs with vanishing error; $E_{\mathrm{D}}^{B\to A}$ is defined with the roles exchanged. We call the two the *directional rates*, and *usable entanglement* means either of them. Both are unchanged by local isometries, and neither increases under local channels of either party or under classical messages in the permitted direction. A local measurement by Alice whose outcome is announced to Bob, succeeding with probability p and leaving the conditional state σ, gives $E_{\mathrm{D}}^{A\to B}(\rho) \ge p\, E_{\mathrm{D}}^{A\to B}(\sigma)$: a fraction p of the copies succeed, Alice names them, and the rest are discarded. The hashing inequality of Devetak and Winter [51] states

$$E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ \ge \ S(\rho_B) - S(\rho_{AB}), \tag{2.2}$$

the coherent information from Alice to Bob, achieved with classical communication from Alice to Bob. We use it in two standard forms. If an N-qubit pair state has been made Bell-diagonal by a bilateral Pauli twirl (Alice samples a Pauli, announces it, and both parties apply it, Bob its complex conjugate), then $S(\rho_B) = N$ and $S(\rho_{AB})$ is the Shannon entropy of the $2N$ Bell-error bits, which is at most the sum of their binary entropies, however correlated they are [14]:

$$E_{\mathrm{D}}^{A\to B} \;\geq\; N - \sum_{k=1}^{2N} h(p_k), \tag{2.3}$$

where p_k is the probability that the k-th Bell stabilizer fails. If instead a state on $\mathbb{C}^D \otimes \mathbb{C}^D$ has singlet fraction $1-p$ with $p < \frac{1}{2}$, an isotropic twirl and (2.2) give

$$E_{\mathrm{D}} \;\geq\; \log D - h(p) - p\log(D^2-1) \;\geq\; (1-2p)\log D - h(p) \tag{2.4}$$

in either direction. For a pure state both directional rates equal the entanglement entropy: this is the concentration theorem of Bennett, Bernstein, Popescu and Schumacher [12].

The resource of a game. For a finite game G with its original question distribution and legal-answer alphabets, the *loss* of a strategy is one minus its winning probability, and ε always denotes a loss. Put

$$d_G^{\rightarrow}(\varepsilon) = \inf\{E_{\mathrm{D}}^{A\to B}(\rho_{AB}) : \text{ some finite-dimensional strategy on } \rho_{AB} \text{ has loss} \leq \varepsilon\}, \tag{2.5}$$

the measurements being arbitrary POVMs, and define $d_G^{\leftarrow}$ with $E_{\mathrm{D}}^{B\to A}$. An infimum need not be attained. Marginal entropy, pure Schmidt rank, ambient dimension, mixed-state E_{F} and mixed-state E_{D} are five different quantities, and each statement below names the one it bounds.

Precision. The letter ε is reserved for the loss of a game from value one. Source statements use a *precision* e, a bound on root errors of relations, measured as vector or Hilbert–Schmidt norms. A game of loss ε supplies source precision $e = C_G\sqrt{\varepsilon}$ (Section 4), and a source model of precision e returns a strategy of loss $O(e^2)$. Neither convention is a relative bias deficit or a gap below a Bell optimum; the latter appears only in Section 16, where it is called a Bell gap.

The conversion of errors. Every rate below passes through the same chain of conversions. A tested incidence lost with probability ℓ gives an actual mirror of error $2\sqrt{\ell}$, so loss ε gives mirror errors $\delta = O(\sqrt{\varepsilon})$ (Theorem 4.3 and Proposition 6.1). Mirrors give state commutators at most $\sqrt{2}\,\delta$ (Lemma 2.11(d)). A word of length L and area A in the source relations then has error $O((A+L)\delta)$ (Theorem 13.5); charging each cell its own conjugator (Lemmas 2.3 and 2.4) gives the simpler $O(A(L+A)\delta)$. Either is $O(\delta)$ for the fixed words of a fixed source, and the cost is charged explicitly where the words grow. For $2N$ extracted Pauli labels whose signed pair errors, state commutators and actual mirror errors are all at most κ, the Bell-stabilizer failures are $p_k \leq 25N\kappa^2$ (Theorem 6.4), and hashing (2.3) turns these into $E_{\mathrm{D}}^{A\to B} \geq N[1 - 2h(25N\kappa^2)]$. In the other direction a model of precision e gives a strategy of loss $O(e^2)$ (Theorem 4.5). Between game loss and source precision the conversion is thus $e = C_G\sqrt{\varepsilon}$, with loss $O(e^2)$ on return. Further power losses, from spectral cuts and from growing decoders, are accounted for in the proofs where they occur. Constant factors in precision change a law $\alpha\log(1/\varepsilon)$ only by $O(1)$ and leave every exponent unchanged.

Original states and actual mirrors. A strategy on a mixed state ρ_{AB} is analysed on a purification $\Psi \in H_A \otimes H_B \otimes H_E$, which is used only in proofs: no operation ever acts on E. If U acts on Bob and V on Alice, we call V an *actual mirror* of U with error δ if $\|(V-U)\Psi\| \le \delta$; since $\|(V-U)\Psi\|^2 = \operatorname{Tr} \rho_{AB}(V-U)^*(V-U)$, this does not depend on the purification. Opposite-party operators commute, and that is the only algebraic fact behind the following transfers.

Lemma 2.11 (Transfer across the parties)**.** *Let Ψ purify ρ_{AB}.*

(a) If $U_1, \dots, U_l$ act on Bob and $V_1, \dots, V_l$ on Alice, all contractions, with $\|(V_i - U_i)\Psi\| \le \delta_i$, then $\|(U_1 \cdots U_l - V_l \cdots V_1)\Psi\| \le \sum_i \delta_i$.

(b) If moreover F acts on Bob and the U_i and V_i are unitaries, then $\|FU_1 \cdots U_l\Psi\| \le \|F\Psi\| + \|F\|_{\mathrm{op}} \sum_i \delta_i$.

(c) If R acts on Bob and S on Alice, both unitary, then $\|(R^n - S^n)\Psi\| \le n\|(R-S)\Psi\|$ for $n \ge 0$.

(d) If U is a unitary on Bob and V a unitary on Alice with $\|(V-U)\Psi\| \le \mu$, then $\|[U, \rho_B^{1/2}]\|_2 \le \sqrt{2}\mu$ and $\|[V, \rho_A^{1/2}]\|_2 \le \sqrt{2}\mu$.

The same holds with the parties exchanged.

Proof. (a) Replace the rightmost factor U_l, which acts first on Ψ, by V_l, at cost δ_l since the left factors are contractions; then move V_l to the far left, past the Bob operators, and induct. (b) By (a), $\|FU_1 \cdots U_l\Psi - FV_l \cdots V_1\Psi\| \le \|F\|_{\mathrm{op}} \sum_i \delta_i$, and $\|FV_l \cdots V_1\Psi\| = \|V_l \cdots V_1 F\Psi\| = \|F\Psi\|$. (c) $R^n - S^n = \sum_{j<n} R^{n-1-j}(R-S)S^j$, and $(R-S)S^j\Psi = S^j(R-S)\Psi$, while $R^{n-1-j}S^j$ is unitary. (d) Write $\xi = \rho_B^{1/2}$ and use a Schmidt decomposition $\Psi = \sum_k \sqrt{\lambda_k}\, e_k \otimes f_k$ across B versus AE, with Π the projection onto the span of the e_k. The isometry $\Theta(e_k \otimes v) = |v\rangle\langle f_k|$ of the standard form (Lemma 2.1) maps Ψ to ξ, $(I \otimes U)x$ to $U\Theta(x)$, and $(\Pi(V \otimes I_E) \otimes I)\Psi$ to ξC, where C is the transposed matrix of the compression $\Pi(V \otimes I_E)\Pi$ in the bases (e_k), (f_k); the part of $(V \otimes I_E)\Psi$ outside the range of Π is orthogonal to $U\Psi$. So C is a contraction on H_B with $\|U\xi - \xi C\|_2 \le \mu$, and the same for the adjoints gives $\|U^*\xi - \xi C^*\|_2 \le \mu$; the adjoints have the same mirror error, because $(V^* - U^*)\Psi = -V^*U^*(V-U)\Psi$ once U^* is commuted past V. With $x = \operatorname{Tr}(\xi U^* \xi U) \ge 0$, $y = \|\xi C\|_2$, $y' = \|\xi C^*\|_2$ and the positive form $\langle K, L\rangle_\xi = \operatorname{Tr}(\xi K^* \xi L)$, which is the Hilbert–Schmidt inner product of $\xi^{1/2}K\xi^{1/2}$ and $\xi^{1/2}L\xi^{1/2}$, expanding $\|U\xi - \xi C\|_2^2 = 1 + y^2 - 2\operatorname{Re}\langle U, C\rangle_\xi$ and applying the Cauchy–Schwarz inequality to the form, with $\langle U, U\rangle_\xi = x$ and $\langle C, C\rangle_\xi \le \|C\xi\|_2 \|\xi C\|_2 = y'y$, gives $\mu^2 \ge 1 + y^2 - 2\sqrt{xyy'}$, and likewise $\mu^2 \ge 1 + y'^2 - 2\sqrt{xyy'}$. Adding, and using $y^2 + y'^2 \ge 2yy'$, gives $2\mu^2 \ge 2 - 2x + 2(\sqrt{yy'} - \sqrt{x})^2 \ge 2 - 2x = \|[U, \xi]\|_2^2$. Exchange the parties for V. □

Part (d) is the unitary form of Lemma 2.2, with the same constant; for reflections the inverse mirror is the mirror itself. Square-root-state commutators are therefore *consequences* of actual mirrors.

The state-weighted source contract (SW). Fix a finite presentation with a marked central involution J. A *model at precision e* is an assignment of unitaries to the generators on a finite-dimensional space and a density matrix ρ, $\xi = \rho^{1/2}$, possibly singular, such that every listed relator R, every generator U and the marked word satisfy

$$\|(R - I)\xi\|_2 \le e, \qquad \|[U, \xi]\|_2 \le e, \qquad \|(J + I)\xi\|_2 \le e,$$

with the unnormalized Hilbert–Schmidt norm and $\operatorname{Tr}\rho = 1$. For the flat state this is the normalized-trace approximate representation with J near $-I$. This is Definition 4.2 below.

The exact-source contract (EX). Bob's operators satisfy all relations of the presentation exactly, as operator identities, in the sector $\pi(J) = -I$, on a possibly infinite-dimensional ambient space; the input is the *original* arbitrary normal mixed state ρ_{AB}; and Alice has actual unitaries mirroring each signed generator with error at most e on a purification. Word errors follow from Lemma 2.11(a), applied to the unchanged vector in reverse order. The exact upper constructions below use finite-rank states inside infinite exact representations; none of the exact sources has a nonzero finite-dimensional representation with $J = -I$. "Both directions" always means the two physical one-way rates $E_{\mathrm{D}}^{A\to B}$ and $E_{\mathrm{D}}^{B\to A}$ of the original state.

I. The engine

3 The qubit count in an arbitrary state

Let $X_1, \dots, X_M, Z_1, \dots, Z_M$ be involutions. If the Z_i commute, X_i anticommutes with Z_i and commutes with Z_j for $j \neq i$, then the joint eigenspaces of the Z_i carry labels $z \in \{\pm 1\}^M$ and X_i maps the label z to the label with its i-th sign flipped, so the dimension is a multiple of 2^M. Our companion paper proves a robust form of this count in an arbitrary state ρ, with every error measured in that state and the errors of different pairs adding in squares [101, Theorem C, Theorem 4.1 and Corollary 4.2]. We state it in the form below; the proof is not repeated. The price of a non-flat state is the state commutators $\|[Z_i, \xi]\|_2$ and $\|[X_i, \xi]\|_2$, which cannot be dropped [101, Example 4.5]; in games they are supplied by consistency (Lemma 2.2) or by actual mirrors (Lemma 2.11(d)).

Theorem 3.1 (Qubit count; [101, Theorem 4.1])**.** *Let $X_1, \dots, X_M, Z_1, \dots, Z_M$ be involutions on $\mathbb{C}^d$, ρ a density matrix on $\mathbb{C}^d$ and $\xi = \rho^{1/2}$. Put $c_i^Z = \|[Z_i, \xi]\|_2$, $c_i^X = \|[X_i, \xi]\|_2$, $f(g) = Z_1^{g_1} Z_2^{g_2} \cdots Z_M^{g_M}$ for $g \in \mathbb{Z}_2^M$, and, for uniformly random g,*

$$a_i^2 = \mathbb{E}_g \|(f(g + e_i) - f(g) Z_i)\xi\|_2^2, \qquad b_i^2 = \mathbb{E}_g \|((-1)^{g_i} X_i f(g) - f(g) X_i)\xi\|_2^2.$$

Then

(i) $a_i^2 \le 2 \sum_{j>i} (\|[Z_i, Z_j]\xi\|_2 + c_j^Z)^2 + 2 \sum_{j>i} (c_j^Z)^2$;

(ii) $b_i^2 \le 2(\|\{X_i, Z_i\}\xi\|_2 + c_i^Z)^2 + 2 \sum_{j\neq i} (\|[X_i, Z_j]\xi\|_2 + c_j^Z)^2 + 2 \sum_{j=1}^M (c_j^Z)^2$;

(iii) $S(\rho) \;\ge\; \sum_{i=1}^M \Big[1 - \dfrac{(b_i + c_i^X)^2}{2 \ln 2} - h\Big(\min\Big(\dfrac{(a_i + c_i^Z)^2}{4}, \dfrac{1}{2}\Big)\Big)\Big]$.

Corollary 3.2 (Uniform errors; [101, Corollary 4.2])**.** *In the setting of Theorem 3.1, let $M \ge 2$ and suppose that, for all $i \neq j$,*

$$\|[Z_i, Z_j]\xi\|_2, \;\; \|[X_i, Z_j]\xi\|_2, \;\; \|\{X_i, Z_i\}\xi\|_2, \;\; c_i^X, \;\; c_i^Z \;\le\; \kappa.$$

If $16 M \kappa^2 \le \frac{1}{10}$, then $S(\rho) \ge \frac{3}{4} M$.

The proof embeds the space coherently into a register indexed by $\{0,1\}^M$ through the products $f(g)$, on which the shifts and signs are exact Pauli operators; the two-sided twirl of Lemma 2.9 computes the intertwining errors a_i, b_i from the pairwise relations, pinching the embedded state gives about one bit per pair, and a Jensen–Shannon bound shows that each pinching raises the entropy only a little [101, Section 4]. The count is applied on the sector where a central involution acts as -1, and the entropy of that sector state controls the entropy of the whole state [101, Lemma 4.9].

Remark 3.3 (Comparison). The stability theorem of Gowers and Hatami [66] needs the defect of an approximate homomorphism of the whole Pauli group, and the many-qubit tests of Natarajan and Vidick [107] certify n EPR pairs from a test that samples the whole Pauli group, through a state-dependent form of that theorem; Pauli stability for synchronous strategies in a tracial von Neumann algebra is in [49], and Kaniewski, Tomamichel and Wehner use state-dependent anticommutators for entropic uncertainty relations [78]. The count uses only the pairwise relations of $2M$ involutions, measured in an arbitrary state, and its errors add in squares. It bounds the entropy of Bob's marginal, which is not usable entanglement (Section 1.6); its operational counterpart on the original mixed state is the signed regular extraction of Theorem 6.4, which uses the same twirl on both parties. The companion paper also applies the count to the CHSH(n) games: the least pure-state entanglement entropy, and the least entanglement of formation, of a strategy within a factor $1-\varepsilon$ of the optimal bias are $\Theta(\min(n, 1/\varepsilon))$ [101, Theorem E], which improves the exponent $\frac{1}{5}$ of Ostrev and Vidick [111] to the optimal one. Whether the near-optimal strategies of CHSH(n) contain $\Omega(\min(n, 1/\varepsilon))$ distillable ebits is open (Section 18).

4 Transfer and return with the state kept

Slofstra and Vidick showed that near-optimal strategies for a linear system game give approximate representations of its solution group, and that the entanglement of the game is controlled by the hyperlinear profile of the central element J [134]. For a general state they reach a representation that is accurate to order $\varepsilon^{1/4}$: winning with probability $1-\varepsilon$ controls the relations, and the commutators of Bob's observables with $\rho^{1/2}$, to order $\sqrt{\varepsilon}$ in the state [134, Proposition 5.4], and the passage to a spectral projection of ρ halves the exponent again [134, Theorem 5.1 and Lemmas 5.5, 5.6]. Our companion paper keeps the state in both directions for linear system games [101, Theorem D]: a strategy gives an assignment satisfying three conditions measured in Bob's reduced state at precision $\sqrt{\varepsilon}$, and conversely such an assignment gives a strategy of loss $O(\varepsilon)$ on a purification of the same state. This section states that correspondence in the form used here and proves its extension to Boolean constraint games (Theorem 4.9), which the universal compiler of Section 5.3 needs, and the comparison of whole spectra (Corollary 4.10).

Definition 4.1 (Linear system games)**.** Let $A \in \mathbb{F}_2^{m\times n}$ and $b \in \mathbb{F}_2^m$, let $V_i = \{j : A_{ij} = 1\}$ be nonempty, and $s = \max_i |V_i|$. The *solution group* $\Gamma(A, b)$ is generated by $x_1, \dots, x_n, J$ with the relators

$$x_j^2, \quad J^2, \quad [x_j, J] \ \ (1 \le j \le n), \qquad [x_j, x_k] \ \ (j, k \in V_i),$$
$$J^{b_i} x_{j_1} x_{j_2} \cdots x_{j_r} \ \ (V_i = \{j_1 < \cdots < j_r\})$$

[34, 33]. In the game $G(A, b)$ a referee draws a pair (i, j) from a distribution π on $\{1, \dots, m\} \times \{1, \dots, n\}$. Alice receives i and answers an assignment to the variables of V_i that satisfies equation i; Bob receives j and answers one bit; they win if $j \notin V_i$ or they give x_j the same value. Two distributions are in use: j uniform in V_i after a uniform i [101], and i and j independent and uniform [134]. Put $\pi_{\min} = \min\{\pi(i, j) : j \in V_i\}$; it is at least $1/(ms)$ for the first and equals $1/(mn)$ for the second.

A strategy consists of a unit vector $\psi \in H_A \otimes H_B$, POVMs $\{P_a^i\}_a$ on H_A indexed by the satisfying assignments of equation i, and POVMs $\{Q_c^j\}_{c=\pm 1}$ on H_B; we write answers as signs. Its observables are $A_{ij} = \sum_a a P_a^i$ for $j \in V_i$ and $B_j = Q_{+1}^j - Q_{-1}^j$, Hermitian contractions, and the pair (i, j) with $j \in V_i$ is won with probability $\frac{1}{2}(1 + \langle \psi, (A_{ij} \otimes B_j)\psi \rangle)$.

Definition 4.2 (State-weighted profile)**.** Let $\langle S \mid R\rangle$ be a finite presentation, $w \in F(S)$ and $e > 0$. The *state-weighted profile* $\mathrm{hlp}^{\rho}(w; e)$ is the least d for which there are a homomorphism $\varphi : F(S) \to \mathrm{U}(d)$ and a density matrix ρ on $\mathbb{C}^d$, with $\xi = \rho^{1/2}$, such that

(S1) $\|(\varphi(r) - I)\xi\|_2 \le e$ for every $r \in R$;

(S2) $\|[\varphi(s), \xi]\|_2 \le e$ for every $s \in S$;

(S3) $\|(\varphi(w) + I)\xi\|_2 \le e$.

We write hlp^{ρ}_{R} when the presentation matters.

The hyperlinear profile $\mathrm{hlp}(w; \delta, \varepsilon)$ of [134, Definition 2.3] is the least dimension of an ε-representation, that is $\|\varphi(r) - I\|_f \le \varepsilon$ for $r \in R$, with $\|\varphi(w) - I\|_f \ge \delta$. For $\delta = 2$ this forces $\varphi(w) = -I$, and the flat state $\rho = I/d$ then satisfies (S1)–(S3) with $e = \varepsilon$. So $\mathrm{hlp}^{\rho}(w; e) \le \mathrm{hlp}(w; 2, e)$: a lower bound on hlp^{ρ} is a stronger statement than the same lower bound on $\mathrm{hlp}(\,\cdot\,; 2, \cdot)$.

Theorem 4.3 (Transfer; [101, Theorem 5.1])**.** *Let $G = G(A, b)$ be a binary linear system game with question distribution π, and let a strategy win G with probability at least $1 - \varepsilon$. Let r be the Schmidt rank of its state and put $\theta = 2\sqrt{\varepsilon/\pi_{\min}}$. There are a homomorphism $\varphi : F(x_1, \dots, x_n, J) \to \mathrm{U}(r)$ and a density matrix ρ of full rank on $\mathbb{C}^r$ such that $\varphi(J) = -I$, every $\varphi(x_j)$ is an involution, and, with $\xi = \rho^{1/2}$,*

(a) $\|(\varphi(R) - I)\xi\|_2 \le \max(s, 4)\,\theta$ *for every relator R of $\Gamma(A, b)$;*

(b) $\|[\varphi(x_j), \xi]\|_2 \le \sqrt{2}\,\theta$ *for every j.*

In particular $r \ge \mathrm{hlp}^{\rho}_{\Gamma}(J; \max(s, 4)\theta)$, and, if $\omega_q(G) = 1$,

$$\mathcal{E}(G, \varepsilon) \;\ge\; \mathrm{hlp}^{\rho}_{\Gamma}\big(J;\; 2\max(s, 4)\sqrt{\varepsilon/\pi_{\min}}\big).$$

The density matrix ρ is Bob's reduced density matrix restricted to its support.

The proof [101, Section 5.1] compresses Bob's measurements to the support of his reduced state, replaces his observables by involutions without lowering the winning probability and dilates Alice's measurements (Step 1 there); a tested pair lost with probability ℓ_{ij} has $\|(A_{ij} \otimes I - I \otimes B_j)\psi\| = 2\sqrt{\ell_{ij}}$ (Step 2); the relations follow by transferring products of Bob's observables to Alice's commuting row observables (Step 3); and Lemmas 2.1 and 2.2 give the state commutators (Step 4).

Remark 4.4. For the flat state, (a) is [134, Proposition 4.4(b)] up to constants. For a general state, [134, Proposition 5.4] proves (a) and (b) with unspecified constants, and then passes to a spectral projection of ρ, which gives $\mathrm{hlp}(J; 2, C\varepsilon^{1/4}) \le \mathcal{E}(G, \varepsilon)$ [134, Corollary 5.2]; Theorem 4.3 stops before that step. For boolean constraint system games, [60, Proposition 3.13] and [60, Lemma 2.9] bound by $O(\sqrt{\varepsilon})$ the defect, in the state of an ε-perfect commuting-operator strategy, which need not be tracial, of every relation with a decomposition of bounded size. A square root is also lost when near-optimal strategies for boolean constraint system games are rounded to strategies on a maximally entangled state [112, Corollary 1.3].

4.1 The return direction

Conversely, a model that satisfies (S1)–(S3) at precision e gives a strategy that loses with probability $O(e^2)$, on a purification of the same state [101, Theorem 5.4]. For the flat state this is the classical direction of the correspondence [134, Proposition 4.4(a)]. The point is

that the state, and with it its spectrum, is kept: the game and the state-weighted profile then determine each other up to constant factors in the precision.

Let ρ be a density matrix on $\mathbb{C}^d$ and $\xi = \rho^{1/2}$. The isometry $\Theta : \mathbb{C}^d \otimes \mathbb{C}^d \to M_d(\mathbb{C})$ with $\Theta(v \otimes w) = wv^T$ satisfies $\Theta((Q \otimes I)x) = \Theta(x)Q^T$ and $\Theta((I \otimes Y)x) = Y\Theta(x)$. The *canonical purification* of ρ is $\psi_\rho = \Theta^{-1}(\xi)$. Bob's reduced state is ρ, Alice's is ρ^T, and we write $\eta = \xi^T$ for its square root. So $\|(C \otimes I)\psi_\rho\| = \|\xi C^T\|_2 = \|C\eta\|_2$ for operators C on Alice's side, and, for an operator B on Bob's side,

$$\|(B^T \otimes I - I \otimes B)\psi_\rho\| = \|\xi B - B\xi\|_2 = \|[B, \xi]\|_2. \tag{4.1}$$

Alice therefore imitates Bob's observable B by measuring B^T, at a cost equal to its state commutator.

Theorem 4.5 (Return; [101, Theorem 5.4])**.** *Let $G = G(A, b)$ be a binary linear system game and $s = \max_i |V_i|$. Let $\varphi : F(x_1, \dots, x_n, J) \to \mathrm{U}(d)$ and a density matrix ρ on $\mathbb{C}^d$ satisfy (S1)–(S3) for the presentation of Definition 4.1 and $w = J$, at precision e. Put $B_j = \mathrm{r}(\varphi(x_j))$. Consider the strategy on ψ_ρ in which Bob, on question j, measures B_j, and Alice, on question i with $V_i = \{j_1 < \cdots < j_r\}$, measures the binary projective measurements $\{\frac{1}{2}(I \pm B_{j_1}^T)\}, \dots, \{\frac{1}{2}(I \pm B_{j_r}^T)\}$ in turn and answers the recorded signs, flipping the last one if their product is not $(-1)^{b_i}$. Every question pair (i, j) with $j \in V_i$ is lost with probability at most $C_s e^2$, where*

$$\beta_l = 3\Big(1 + \frac{(l-1)(l-2)}{2}\Big) + \frac{13(l-1)}{\sqrt{2}}, \qquad C_s = \frac{1}{4}\Big[\beta_s^2 + \Big(s^2 + 2 + \sum_{l=1}^{s} \beta_l\Big)^2\Big].$$

The strategy has local dimensions d, and Bob's reduced state is ρ.

The proof [101, Section 5.2] rounds the generators by Lemma 2.5, bounds the commutators and the parity of each equation in the state, and controls Alice's sequential records through the branch commutators $[K_a, \rho_A^{1/2}]$ of her Kraus operators [101, Lemma 5.3].

For $s = 3$ the constant is $C_3 \approx 788.2$; it is far from optimal. Nothing is postselected: every run of Alice's procedure produces a valid answer, and the shared state is ψ_ρ itself.

Corollary 4.6 (Two-sided correspondence; [101, Corollary 5.5])**.** *Let $G = G(A, b)$ be a binary linear system game with $\omega_q(G) = 1$ and question distribution π, put $C_G = 2\max(s, 4)/\sqrt{\pi_{\min}}$, and let C_s be as in Theorem 4.5.*

(a) For every $\varepsilon > 0$, $\mathrm{hlp}_\Gamma^\rho(J; C_G\sqrt{\varepsilon}) \le \mathcal{E}(G, \varepsilon) \le \mathrm{hlp}_\Gamma^\rho(J; \sqrt{\varepsilon/C_s})$.

(b) Let $h_\Gamma^S(e)$ be the infimum of $S(\rho)$ over all models satisfying (S1)–(S3) for $w = J$ at precision e, in any dimension, and $q(\varepsilon)$ the infimum of the entanglement entropy of pure-state strategies that win with probability at least $1 - \varepsilon$. Then $h_\Gamma^S(C_G\sqrt{\varepsilon}) \le q(\varepsilon) \le h_\Gamma^S(\sqrt{\varepsilon/C_s})$.

(c) Let $e_\Gamma(d)$ be the infimum of the precisions e with $\mathrm{hlp}_\Gamma^\rho(J; e) \le d$, and $\ell_G(d)$ the infimum of the losing probabilities of strategies with $H_A = H_B = \mathbb{C}^d$. Then $e_\Gamma(d)^2/C_G^2 \le \ell_G(d) \le C_s\, e_\Gamma(d)^2$ for every d.

The state commutators in (S2) are needed for the return direction, not only for counting [101, Example 5.6]. That flat approximate representations give strategies with quadratically small loss on a maximally entangled state is due to Slofstra and Vidick [134, Proposition 4.4(a)], following [132, Proposition 3.4], and Paddock develops the correspondence between near-optimal strategies and approximate representations, including rounding to maximally entangled strategies [112]. Since both directions keep an arbitrary state, the dimension, the entanglement entropy, or any other function of the spectrum that does not change when zero eigenvalues are added can be read off the state-weighted profile (Corollary 4.10).

4.2 Boolean constraint games

The same correspondence holds for games whose constraints are arbitrary boolean predicates. We need it because the universal compiler of Section 5.3 produces such games.

Definition 4.7 (Boolean constraint games)**.** A *boolean constraint system* B has variables $x_1, \dots, x_n$ and contexts $C_1, \dots, C_m$, each an ordered list $(j_1, \dots, j_r)$ of distinct variables with a nonempty set $\mathcal{S}_i \subseteq \{\pm 1\}^r$ of allowed assignments; put $k = \max_i |C_i|$. In its game a referee draws (i, j) with $j \in C_i$ from a distribution π; Alice receives i and answers an allowed assignment, Bob receives j and answers a sign, and they win if Alice's assignment gives x_j Bob's sign. Put $\pi_{\min} = \min\{\pi(i,j) : j \in C_i\}$. A linear system game is the case of parity predicates.

Definition 4.8 (Boolean profile)**.** For a context $C_i = (j_1, \dots, j_r)$ and $\alpha \in \{\pm 1\}^r$ let $R_\alpha = \prod_{t=1}^{r} \frac{1}{2}(I + \alpha_t B_{j_t})$, the product taken in the order of the context. A *model* of B at precision e is a family of involutions B_j on $\mathbb{C}^d$ and a density matrix ρ, $\xi = \rho^{1/2}$, with

(B1) $\|[B_j, \xi]\|_2 \le e$ for every j;

(B2) $\|[B_j, B_l]\xi\|_2 \le e$ for j, l in a common context;

(B3) $\|R_\alpha \xi\|_2 \le e$ for every context and every assignment α that it forbids.

Let $h_B(e)$ be the least dimension of a model, $h_B^S(e)$ the infimum of $S(\rho)$, and $\Lambda_B(e)$ the set of nonzero spectra of ρ.

Theorem 4.9 (Boolean correspondence)**.** *Let B be a boolean constraint system with question distribution π, and put $\theta = 2\sqrt{\varepsilon/\pi_{\min}}$,*

$$A_k = \frac{k}{2} + \frac{(k-1)(k+2)}{8}, \qquad K_k = 2^k\Big[A_k^2 + \Big(1 + \frac{k}{2}\Big)^2\Big].$$

(a) A strategy that wins with probability at least $1 - \varepsilon$ and has Schmidt rank r gives a model of B on $\mathbb{C}^r$ at precision $\max(4, k/2)\,\theta$, whose density matrix is Bob's reduced density matrix restricted to its support.

(b) A model of B at precision e on $\mathbb{C}^d$ gives a strategy on the canonical purification of its state, with local dimensions d and Bob's reduced state ρ, that loses every question pair with probability at most $K_k e^2$.

Consequently, if $\omega_q = 1$, then $h_B(\max(4, k/2)\theta) \le \mathcal{E}(G, \varepsilon) \le h_B(\sqrt{\varepsilon/K_k})$, and the same holds for the least entanglement entropy of a pure strategy, with h_B^S in place of h_B. For $k \le 3$, $K_k \le K_3 = \frac{221}{2}$.

Proof. (a) The reductions of the proof of Theorem 4.3 [101, proof of Theorem 5.1, Step 1] apply verbatim: Bob's measurements are compressed to the support of his reduced state, replaced by involutions B_j without lowering the winning probability, and Alice's are dilated, which does not change Bob's state. For each context i, Alice's observables A_{ij}, $j \in C_i$, then commute and their joint spectral projections onto forbidden assignments vanish. A pair (i, j) is lost with probability at most $\varepsilon/\pi_{\min}$, so $\|(A_{ij} \otimes I - I \otimes B_j)\psi\| \le \theta$. Lemmas 2.1 and 2.2 give $\|[B_j, \xi]\|_2 \le \sqrt{2}\,\theta$, which is (B1). For (B2), replacing B_l and then B_j by A_{il} and A_{ij} on ψ costs 2θ in each order, and A_{ij}, A_{il} commute, so $\|[B_j, B_l]\xi\|_2 = \|(I \otimes [B_j, B_l])\psi\| \le 4\theta$. For (B3), replace the factors $\frac{1}{2}(I + \alpha_t B_{j_t})$ of R_α, from the right, by Alice's projections $\frac{1}{2}(I + \alpha_t A_{ij_t})$; each replacement costs $\theta/2$, since the factors are contractions and operators of the two players commute. The result is Alice's joint spectral projection onto α, which is zero. So $\|R_\alpha \xi\|_2 \le k\theta/2$.

(b) Let Bob measure B_j, and let Alice, on question i with $C_i = (j_1, \dots, j_r)$, measure $B_{j_1}^T, \dots, B_{j_r}^T$ in turn on ψ_ρ (Section 4.1) and record the signs α. Her Kraus operator for α is $\frac{1}{2}(I + \alpha_r B_{j_r}^T) \cdots \frac{1}{2}(I + \alpha_1 B_{j_1}^T) = R_\alpha^T$, so through the isometry Θ the branch in which Bob's sign is b is $P_j^b \xi R_\alpha$, with $P_j^b = \frac{1}{2}(I + bB_j)$. Fix $j = j_s$ and α, and put $Q = P_{j_1}^{\alpha_1} \cdots P_{j_{s-1}}^{\alpha_{s-1}}$. Since $P_j^{-\alpha_s} P_j^{\alpha_s} = 0$,

$$P_j^{-\alpha_s} \xi R_\alpha = P_j^{-\alpha_s} [\xi, R_\alpha] + [P_j^{-\alpha_s}, Q] P_{j_s}^{\alpha_s} \cdots P_{j_r}^{\alpha_r} \xi.$$

The first term has norm at most $\frac{1}{2} \sum_t \|[B_{j_t}, \xi]\|_2 \le ke/2$. Expanding the commutator over the factors of Q, the term with $[P_j^{-\alpha_s}, P_{j_l}^{\alpha_l}] = -\frac{1}{4}\alpha_s \alpha_l [B_j, B_{j_l}]$ is followed by $r - l$ contractions whose state commutators add up to at most $(r-l)e/2$; moving them past ξ costs $(r-l)e/4$, so the term is at most $(r-l+1)e/4$. Summing over $l < s$ bounds the branch by $A_k e$. Over all 2^r raw outcomes, the probability that Alice's record disagrees with Bob's is at most $2^k A_k^2 e^2$. A forbidden raw outcome has probability $\|\xi R_\alpha\|_2^2 \le (\|R_\alpha \xi\|_2 + \|[R_\alpha, \xi]\|_2)^2 \le (1 + k/2)^2 e^2$, by (B3). Alice replaces a forbidden outcome by a fixed allowed assignment; this changes her answer only on that event, so every pair is lost with probability at most $K_k e^2$. Nothing is postselected and the shared state is ψ_ρ.

The consequences follow as in Corollary 4.6, and $K_3 = 8[(11/4)^2 + (5/2)^2] = 221/2$. □

For a linear system game the algebras of Definition 4.8 and of the solution group are the same (Example 5.1), so the profiles h_B and $\mathrm{hlp}_\Gamma^\rho(J; \cdot)$ agree up to constant factors in the precision (Proposition 5.3 and Lemma 5.4); Theorems 4.3 and 4.5 give explicit constants for that case. The transfer direction is close to [60, Proposition 3.13 and Lemma 2.9] (Remark 4.4); sequential measurement of commuting observables is standard, and what the return direction adds is that the state is kept.

4.3 The whole spectrum

Both directions keep the state, so they compare whole spectra. For a game G let $\Lambda_G(\varepsilon)$ be the set of nonzero spectra of Bob's reduced density matrices in pure strategies that win with probability at least $1 - \varepsilon$.

Corollary 4.10 (Spectral correspondence)**.** *For a linear system game with solution group Γ, in the setting of Corollary 4.6,*

$$\Lambda_G(\varepsilon) \subseteq \Lambda_\Gamma(C_G \sqrt{\varepsilon}), \qquad \Lambda_\Gamma(e) \subseteq \Lambda_G(C_s e^2),$$

with Λ_Γ as in Section 5.1; for a boolean constraint game the same holds with $\Lambda_B(\max(4, k/2)\theta)$ and $\Lambda_B(e) \subseteq \Lambda_G(K_k e^2)$. In particular, for every function Φ of finite lists of positive numbers, such as a Rényi entropy of any order, the logarithm of the length, or the largest entry, the infimum of Φ over pure strategies at losing probability ε lies between its infima over models at precisions $C_G\sqrt{\varepsilon}$ and $\sqrt{\varepsilon/C_s}$.

Proof. Theorem 4.3 uses Bob's reduced state on its support, whose nonzero spectrum is that of Bob's state; Theorem 4.5 uses the canonical purification, whose reduced state on Bob's side is ρ. The boolean case is Theorem 4.9. □

The Rényi entropy of order 0 here is the logarithm of the rank of ρ, not of the ambient dimension; the two differ when ρ is singular, and the dimension statements of Corollary 4.6 are separate. Combined with Propositions 5.3 and 5.5, the inclusions transfer through the compilers of Section 5: a corner map keeps the spectrum, and an upper map with N blocks replaces it by D_N of it.

Remark 4.11 (Parity models as Boolean models, and large arity). A linear system is a boolean constraint system whose forbidden assignments are those of the wrong parity. A model of the solution group at precision e gives involutions satisfying (B1)–(B3) at precision $D_k e$, with $D_k = \max\{1, [1 + k(k-1)(3k+1)/4]/2\}$ for contexts of size at most k, and then the boolean return loses at most $K_k D_k^2 e^2$. The constant K_k grows quickly: $K_3 = \frac{221}{2} < 111$, but $K_9 = 138496$ and $K_9 D_9^2 = 8829985600$. The arity-three constant must not be used for games of larger arity.

5 Compiling groups into games

Section 4 identifies the cost of a constraint game with the state-weighted profile of an algebra attached to it. To give a game a prescribed cost, one needs a game whose algebra has the profile of a given group. This section supplies the tools, all of which keep the state: algebraic maps between finitely presented algebras of unitaries (Section 5.1), corner maps that recover a group inside a game algebra on the whole space (Section 5.2), a compiler of every finitely presented group with a central involution into a fixed boolean constraint game (Section 5.3), and three routes into linear system games (Section 5.4).

5.1 Presented algebras and algebraic maps

A *presented algebra* $\mathcal{A} = \langle X \mid R \rangle$ is a finite set X and a finite set R of elements of the group algebra $\mathbb{C}[F(X)]$ of the free group. Let I be the two-sided ideal generated by $R \cup R^*$. The algebra is the unital $*$-algebra $\mathbb{C}[F(X)]/I$, in which the generators are unitaries; it is an algebraic quotient, not a completion. A *model* of $\mathcal{A}$ at precision e on $\mathbb{C}^d$ is a homomorphism $\varphi : F(X) \to \mathrm{U}(d)$ and a density matrix ρ, with $\xi = \rho^{1/2}$, such that

$$\|P(\varphi)\xi\|_2 \le e \quad (P \in R), \qquad \|[\varphi(x), \xi]\|_2 \le e \quad (x \in X), \tag{5.1}$$

where $P(\varphi) = \sum_u c_u \varphi(u)$ for $P = \sum_u c_u u$. Write $h_{\mathcal{A}}(e)$ for the least d, $h^S_{\mathcal{A}}(e)$ for the infimum of $S(\rho)$, and $\Lambda_{\mathcal{A}}(e)$ for the set of nonzero spectra of the density matrices of such models, each spectrum a finite list of positive numbers.

Example 5.1. (a) For a finite presentation $\langle S \mid R \rangle$ and a word w, the *negative quotient* $A^- = \langle S \mid \{r-1 : r \in R\} \cup \{w+1\} \rangle$ is $\mathbb{C}[G]/\langle w+1 \rangle$, and its models are those of Definition 4.2; so $h_{A^-}(e) = \mathrm{hlp}^{\rho}(w; e)$. We write h_G, h^S_G and Λ_G for its profiles. For a solution group, $w = J$, and A^-_Γ is the negative quotient of [34, 114].

(b) For a boolean constraint system B (Definition 4.7), the presented algebra $\mathcal{A}(B)$ has a generator for each variable and the relators $x^2 - 1$, $xy - yx$ for variables in a common context, and the ordered products R_α of Definition 4.8 for every forbidden assignment α. It is the algebra of the game [114]. For a linear system (A, b) it is the negative quotient of the solution group, since in it a parity context forces its product to be ± 1.

Lemma 5.2 (Finite identities). *Let (φ, ρ) be a model of $\langle X \mid R \rangle$ at precision e. Every $p = \sum_u c_u u \in \mathbb{C}[F(X)]$ has $\|p(\varphi)\|_{\mathrm{op}} \le M_p$ and $\|[p(\varphi), \xi]\|_2 \le L_p e$, with $M_p = \sum_u |c_u|$ and $L_p = \sum_u |c_u||u|$. Every $P \in I$ has $\|P(\varphi)\xi\|_2 \le C_P e$, where C_P depends only on a chosen expression of P as a finite sum of terms aQb with $a, b \in \mathbb{C}[F(X)]$ and $Q \in R \cup R^*$.*

Proof. The first two bounds follow from the triangle inequality and Lemma 2.3(a). For $Q \in R$, $\|Q(\varphi)\xi\|_2 \le e$, and for $Q = P^*$ with $P \in R$, $\|Q(\varphi)\xi\|_2 = \|\xi P(\varphi)\|_2 \le \|P(\varphi)\xi\|_2 + \|[P(\varphi), \xi]\|_2 \le (1 + L_P)e$. For a term aQb write $b(\varphi)\xi = \xi b(\varphi) + [b(\varphi), \xi]$; the term is at most $M_a(M_b\|Q(\varphi)\xi\|_2 + M_Q L_b e)$. Sum over the terms. □

An element of I is a polynomial for which such an expression exists; the expression is a finite certificate, and the constants below depend on the chosen certificates, not on d or ρ. For $N \geq 1$ let D_N replace each entry λ of a spectrum by N entries λ/N, the spectrum of $I_N/N \otimes \rho$.

Proposition 5.3 (Algebraic maps transfer profiles)**.** *Let $\mathcal{A}_1 = \langle X_1 \mid R_1 \rangle$ and $\mathcal{A}_2 = \langle X_2 \mid R_2 \rangle$ be presented algebras, $N \geq 1$, and let $A_x \in M_N(\mathbb{C}[F(X_1)])$, $x \in X_2$, be matrices such that $x \mapsto A_x$, $x^{-1} \mapsto A_x^*$ induces a unital $*$-homomorphism $\mathcal{A}_2 \to M_N(\mathcal{A}_1)$. There is a constant C such that every model of $\mathcal{A}_1$ at precision e on $\mathbb{C}^d$ gives a model of $\mathcal{A}_2$ at precision Ce on $\mathbb{C}^{Nd}$ with density matrix exactly $I_N/N \otimes \rho$. Hence*

$$h_{\mathcal{A}_2}(Ce) \leq N\, h_{\mathcal{A}_1}(e), \qquad h^S_{\mathcal{A}_2}(Ce) \leq h^S_{\mathcal{A}_1}(e) + \log N, \qquad D_N(\Lambda_{\mathcal{A}_1}(e)) \subseteq \Lambda_{\mathcal{A}_2}(Ce).$$

Proof. Put $\xi_2 = I_N/\sqrt{N} \otimes \xi$ and $\hat{A}_x = A_x(\varphi)$, evaluated entrywise. If every entry of $Y \in M_N(\mathbb{C}[F(X_1)])$ lies in I_1, then $\|Y(\varphi)\xi_2\|_2^2 = \frac{1}{N}\sum_{i,j}\|Y_{ij}(\varphi)\xi\|_2^2 = O(e^2)$ by Lemma 5.2; likewise $\|[Y(\varphi), \xi_2]\|_2 = O(e)$ and $\|Y(\varphi)\|_{\mathrm{op}} = O(1)$ for every Y. Since the class of A_x is unitary in $M_N(\mathcal{A}_1)$, the entries of $A_x^*A_x - 1$ and $A_xA_x^* - 1$ lie in I_1, and Lemma 2.7 gives unitaries U_x with $\|(U_x - \hat{A}_x)\xi_2\|_2$, $\|\xi_2(U_x - \hat{A}_x)\|_2$ and $\|[U_x, \xi_2]\|_2$ all $O(e)$. Let $P \in R_2$, and let $P(U)$ and $P(\hat{A})$ be its values with $x^{\pm 1}$ replaced by $U_x^{\pm 1}$, respectively by $\hat{A}_x$ and $\hat{A}_x^*$. The entries of the image of P lie in I_1, so $\|P(\hat{A})\xi_2\|_2 = O(e)$. For each word $y_1 \cdots y_m$ of P, with letters $Y_l = U^{\pm 1}$ and Y_l' the corresponding $\hat{A}$ or $\hat{A}^*$,

$$Y_1 \cdots Y_m - Y_1' \cdots Y_m' = \sum_l Y_1' \cdots Y_{l-1}'(Y_l - Y_l')R_l, \qquad R_l = Y_{l+1} \cdots Y_m,$$

and $(Y_l - Y_l')R_l\xi_2 = (Y_l - Y_l')\xi_2 R_l + (Y_l - Y_l')[R_l, \xi_2]$ with R_l unitary. The first term is $O(e)$, using $\|(U^* - \hat{A}^*)\xi_2\|_2 = \|\xi_2(U - \hat{A})\|_2$ for inverted letters, and so is the second. So $\|P(U)\xi_2\|_2 = O(e)$ for every $P \in R_2$, and every U_x has state commutator $O(e)$. The dimension is Nd, the state is $I_N/N \otimes \rho$, and $S(I_N/N \otimes \rho) = S(\rho) + \log N$. □

The identity maps between two finite presentations of the same algebra satisfy the hypothesis with $N = 1$, so profiles depend on the presentation only through the constant in the precision. For instance a Tietze transformation of a group presentation, or the passage between a solution group and the algebra of its linear system game, changes e by a constant factor and keeps the state.

Lemma 5.4 (Rounding to involutions)**.** *Let B be a boolean constraint system with contexts of size at most k. A model of $\mathcal{A}(B)$ at precision e gives, by $B_x = \mathrm{r}(\varphi(x))$, involutions satisfying (B1)–(B3) of Definition 4.8 with the same ρ at precision $C_k e$; conversely the involutions of Definition 4.8 at precision e form a model of $\mathcal{A}(B)$ at precision e.*

Proof. The converse is immediate, since $B_x^2 = I$. For the first part, Lemma 2.5 gives $\|(B_x - \varphi(x))\xi\|_2 \leq e$ and $\|[B_x, \xi]\|_2 \leq 3e$. Each relator of $\mathcal{A}(B)$ is a fixed polynomial of degree at most k in contractions, and replacing the letters $\varphi(x)$ by B_x one at a time, as in Lemma 2.3(e), costs $O(e)$ in the state. □

Polar rounding and the replacement of a presentation by another are standard in the theory of approximate representations [134]. What Proposition 5.3 adds is that both are done in one fixed state at linear precision, that the images A_x need be unitary only modulo the relations, which is why the two-sided estimate of Lemma 2.7 is needed, and that the whole spectrum is carried along.

5.2 Corner maps

The lower direction of a comparison needs a map from the group into the game algebra. Restricting a homomorphism keeps the state but does not, by itself, send the distinguished involution to -1. The following maps do, by covering the whole space with corners on which it is -1.

Proposition 5.5 (Corner maps)**.** *Let $G = \langle S_G \mid R_G\rangle$ and $H = \langle S_H \mid R_H\rangle$ be finitely presented, with w_G representing a central involution of G and J one of H. Let $\psi_1, \dots, \psi_k : F(S_G) \to F(S_H)$ induce homomorphisms $G \to H$, put $W_i = \psi_i(w_G)$, and suppose that, in H, the W_i commute with each other and with every $\psi_j(s)$, $s \in S_G$, $1 \le i, j \le k$, and that some $t \in F(S_H)$ satisfies $tbt^{-1} = Jb$ for $b = W_1 \cdots W_k$. In A_H^- put*

$$P = \tfrac{1}{2}(1-b), \qquad P_i = \tfrac{1}{2}(1-W_i), \qquad Q_i = PP_i \prod_{j<i}(1-P_j), \qquad H_s = \sum_i Q_i \psi_i(s) Q_i,$$

and $F(s) = H_s + tH_s t^{-1}$ for $s \in S_G$. Then $s \mapsto F(s)$ defines a unital $*$-homomorphism $A_G^- \to A_H^-$ with $F(w_G) = -1$. Consequently there is a constant C such that every model of H at precision e gives a model of G at precision Ce on the same space with the same density matrix, and

$$h_G(Ce) \le h_H(e), \qquad h_G^S(Ce) \le h_H^S(e), \qquad \Lambda_H(e) \subseteq \Lambda_G(Ce).$$

For $k = 1$ the hypothesis reads $[\psi_1(w_G), t] = J$, and $F(s) = P\psi_1(s)P + tP\psi_1(s)Pt^{-1}$.

Proof. Work in A_H^-, where $J = -1$. The W_i are commuting self-adjoint unitaries, so the P_i are commuting projections, and P and the Q_i are projections. The Q_i are mutually orthogonal, since for $i < j$ the factor $1 - P_i$ of Q_j annihilates P_i, and $\sum_i Q_i = P(1 - \prod_i (1 - P_i)) = P$, because $b\prod_i(1-P_i) = \prod_i W_i \frac{1+W_i}{2} = \prod_i (1-P_i)$ gives $P\prod_i(1-P_i) = 0$. Every Q_i is a polynomial in the W_j, so it commutes with every $\psi_l(s)$. Hence, for words u, v and with H_u defined by the same formula, $H_u H_v = H_{uv}$: the cross terms contain $Q_i Q_j = 0$. Also $H_u^* = H_{u^{-1}}$, $H_r = P$ for $r \in R_G$ because $\psi_i(r) = 1$ in H, and $H_{w_G} = \sum_i Q_i W_i Q_i = -P$ because $W_i Q_i = (1 - 2P_i)Q_i = -Q_i$. Next $tPt^{-1} = \frac{1}{2}(1 - Jb) = 1 - P$, so $H_u = PH_uP$ and $tH_ut^{-1} = (1-P)tH_ut^{-1}(1-P)$, and the cross terms of $F(u)F(v)$ vanish. So $F(u)F(v) = F(uv)$, $F(u)^* = F(u^{-1})$, $F(r) = P + (1-P) = 1$ for $r \in R_G$, and $F(w_G) = -1$. The images are finite polynomials, and Proposition 5.3 with $N = 1$ gives the rest. For $k = 1$, $Q_1 = P$, and $[W, t] = WtW^{-1}t^{-1} = J$ is equivalent to $tWt^{-1} = JW$ for an involution W. □

For the flat profile, the lower bound of [134, Proposition 6.2] uses a witness t with $k = 1$ to show that the image of w_G is far from I [134, Lemma 6.3], and then restricts the representation to the -1 eigenspace of that image [134, Proposition 2.7]; the restriction keeps a flat state flat but cuts out a sector. Proposition 5.5 replaces the restriction by an algebraic map on the whole space: no sector is cut out, and no sector weight multiplies the entropy. A group homomorphism cannot do the same here. If $\psi_1(w_G)$ were identified with J, the relation $[\psi_1(w_G), t] = J$ would become $[J, t] = J$, which forces $J = 1$ since J is central. The case $k > 1$ is used for the extended presentations of Proposition 5.10 and for the restricted linear compiler of Section 11.2, where several conjugates of the involution are needed to cover the space.

5.3 A universal compiler into boolean constraint games

The next theorem compiles every finitely presented group with a central involution into a fixed boolean constraint game whose constraints are parity equations and one binary predicate, at a fixed cost in dimension and entropy.

Theorem 5.6 (Universal compiler)**.** *Let $G = \langle S \mid R\rangle$ be a finite presentation with a generator $w \in S$ representing a central involution. There is a boolean constraint system B_G, computable from the presentation, in which every context has at most three variables and every constraint is either a parity equation or a binary constraint forbidding the assignment $(-1,-1)$, and there are unital $*$-homomorphisms*

$$F : A_G^- \to \mathcal{A}(B_G), \qquad \sigma : \mathcal{A}(B_G) \to M_{12}(A_G^-), \qquad \sigma(F(a)) = I_{12} \otimes a,$$

given by finitely many polynomials. Consequently there are constants $C, C' > 0$ with

$$h_G(Ce) \le h_{B_G}(e) \le 12\, h_G(e/C'), \qquad h_G^S(Ce) \le h_{B_G}^S(e) \le h_G^S(e/C') + \log 12,$$

and $\Lambda_{B_G}(e) \subseteq \Lambda_G(Ce)$, $D_{12}(\Lambda_G(e/C')) \subseteq \Lambda_{B_G}(e)$. If $h_G(e) < \infty$ for every $e > 0$, the game of B_G, with a context chosen uniformly and then a variable of it uniformly, has quantum value one, and there are $C_1, C_2 > 0$ with

$$h_G(C_1\sqrt{\varepsilon}) \le \mathcal{E}(G(B_G), \varepsilon) \le 12\, h_G(C_2\sqrt{\varepsilon}) \qquad (\varepsilon > 0),$$

and the same with h_G^S and $\log 12$ added on the right for the least entanglement entropy of pure strategies.

The first stage. Let B^0 have involutive variables e_1, e_2, e_3, A, T and V_s for $s \in S$, and the following conditions; write $P_i = \frac{1}{2}(1 - e_i)$ and $N = \{A\} \cup \{V_s : s \in S\}$.

(F1) One-hot frame: e_1, e_2, e_3 form a context whose allowed assignments are $(-,+,+)$, $(+,-,+)$, $(+,+,-)$. In the algebra the P_i are orthogonal projections with $P_1 + P_2 + P_3 = 1$.

(F2) Coordinate swaps: every $n \in N$ satisfies $ne_1n = e_2$, $ne_2n = e_1$, $ne_3n = e_3$, and $Te_1T = e_3$, $Te_3T = e_1$, $Te_2T = e_2$.

(F3) Blank support: for every $n \in N$, (e_3, n) is a context forbidding just $(-1,-1)$. In the algebra this says $P_3(1-n) = 0$, so $nP_3 = P_3$.

(F4) Relators: put $U_s = V_sA$. For every $r \in R$, the word $r(U)$, in which s is replaced by V_sA and s^{-1} by AV_s, conjugates T to itself. This is imposed by a chain of conjugacy relations: if $r(U) = g_1 \cdots g_n$ with letters $g_l \in N$, new involutions $t_n = g_nTg_n$, $t_l = g_lt_{l+1}g_l$ for $1 < l < n$, and the relation $g_1t_2g_1 = T$.

(F5) Sign: (V_w, A, e_3) is a context with $V_wAe_3 = -1$.

Upper map. On three blocks over A_G^- let e_1, e_2, e_3 be the diagonal sign matrices with -1 in block $1, 2, 3$, let A exchange blocks 1 and 2 and T blocks 1 and 3, and put

$$V_s = \begin{pmatrix} 0 & s & 0 \\ s^* & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad U_s = V_sA = \operatorname{diag}(s, s^*, 1).$$

For a word r, $r(U) = \operatorname{diag}(r(s), \tilde{r}(s), 1)$, where $\tilde{r}$ replaces each letter by its inverse without reversing the order. If $r = 1$ in G, this commutes with T *whatever $\tilde{r}$ is*; the chain variables are mapped to the corresponding conjugates of T, which are involutions. Finally $w = -1$ gives $U_w = -e_3$, which is (F5). All conditions hold, so this is a unital $*$-homomorphism $\sigma_1 : \mathcal{A}(B^0) \to M_3(A_G^-)$. The middle block need not satisfy any relator, and it is not tested: the swap T exchanges the first block with the third, which is fixed pointwise.

Lower map. In $\mathcal{A}(B^0)$, A and every V_s exchange P_1 and P_2 and fix P_3, so U_s commutes with the P_i, and (F3) gives $U_sP_3 = P_3$; the same holds for U_s^{-1} and every word $r(U)$. By

(F4) and $TP_3T = P_1$, $r(U)P_1 = Tr(U)P_3T = P_1$ for $r \in R$. So $\theta(s) = P_1U_sP_1$ is a unitary representation of G in the corner with unit P_1, and (F5) gives $\theta(w) = -P_1$, since $e_3P_1 = P_1$. Put

$$F_1(s) = \theta(s) + A\theta(s)A + T\theta(s)T.$$

The three terms live in the corners P_1, P_2, P_3, their cross products vanish and their units add up to 1, so F_1 is a unital $*$-homomorphism $A_G^- \to \mathcal{A}(B^0)$ with $F_1(w) = -1$. Evaluated on σ_1 it gives $\sigma_1(F_1(s)) = I_3 \otimes s$.

Second stage. Apply the nice-conjugacy conversion and the six-equation gadget of Slofstra [132, Lemma 4.4 and Proposition 4.2], in the explicit form of [101, proof of Lemma B.1, Steps 3 and 4], to the conjugacy relations of B^0, and keep the contexts (F1), (F3) and (F5) unchanged. The result is a boolean constraint system B_G. The old variables generate a copy of $\mathcal{A}(B^0)$ in $\mathcal{A}(B_G)$, because the new equations imply the old conjugacy relations by finite multiplication [132, 101], and the lifts of the two stages give a unital $*$-homomorphism $\sigma_g : \mathcal{A}(B_G) \to M_4(\mathcal{A}(B^0))$ that sends every old variable x to $I_4 \otimes x$; it respects the kept contexts because these involve only old variables. Finally replace the one-hot context by the parity equation $e_1e_2e_3 = -1$ and three binary constraints forbidding $(-,-)$ on the pairs: among commuting signs the parity equation allows an odd number of negative signs and the binary constraints at most one, so the allowed assignments, and the algebra, are unchanged.

Proof of Theorem 5.6. Put $F = F_1$, followed by the inclusion of $\mathcal{A}(B^0)$, and $\sigma = (\mathrm{id}_4 \otimes \sigma_1) \circ \sigma_g$. Then $\sigma(F(s)) = I_4 \otimes I_3 \otimes s$. Every context of B_G has at most three variables: the gadget equations and (F5) have three, the binary constraints two. The profile bounds follow from Proposition 5.3, applied to F with $N = 1$ and to σ with $N = 12$, together with Lemma 5.4. The game bounds follow from these and Theorem 4.9; if $h_G(e) < \infty$ for every e, the return direction gives strategies whose losing probability tends to 0. □

Two clauses are essential. Without the one-hot frame, $e_1 = e_2 = e_3 = -1$ and $A = T = V_s = 1$ satisfy every other first-stage condition and encode nothing. Without the blank support of one V_s, the source relators need hold only up to a sign in the payload block: for $G = \langle u, v, q, w \mid u^3,\ uvq^{-1},\ w^2,\ [w,u], [w,v], [w,q] \rangle$, the assignment $U_u = -I$, $U_v = U_q = I$, $U_w = \mathrm{diag}(-1,-1,1)$ in three blocks, with $V_s = U_sA$, satisfies every other first-stage condition although u^3 and uvq^{-1} are mapped to $-I$.

Remark 5.7 (Sources and what is new). The parts of this construction are known, with one exception.

- A map in the direction of F exists in much greater generality: Slofstra's embedding theorem gives $A_G^- \hookrightarrow A_\Gamma^-$ into a solution group [133, Theorem 3.1], and Fanizza, Kroell, Mehta, Paddock, Rochette, Slofstra and Zhao embed the group algebra of a group generated by involutions into the algebra of a boolean constraint system, with relation decompositions of bounded size [60, Lemma 3.12]; with [60, Proposition 3.13 and Lemma 2.9] this controls near-perfect strategies in non-tracial states. The relation-decomposition bound used there, [60, Lemma 2.9], is Lemma 3.9 of Mehta, Slofstra and Zhao [99].
- The second stage is Slofstra's [132, Lemma 4.4, Proposition 4.2 and Remark 4.5]; its lift of every old variable to $I_4 \otimes x$ is explicit in [60, Lemma 5.8 and Theorem 5.5], which also gives decomposition bounds and the extension of traces. The chains of (F4) are their flattened nested conjugacies [60, Definition 5.6 and Lemma 5.7].
- The dilation $y \mapsto \mathrm{adiag}(y, y^*)$ behind V_s is that of [132, Proposition 4.8]; the constraint language, one-hot contexts and binary exclusions, is that of Paddock's encoding of synchronous games [112], recalled in [114, Section 3.2]; and the passage from one-hot to parity and exclusions is an instance of [114, Lemma 5.4].

- For synchronous games, Harris's reduction to graph three-colouring [67, Lemma 4.4 and Theorems 4.5 and 4.6] has a whole-algebra split with no amplification: his six colour-permuted corners are orthogonal, central and add up to one. It maps games to games.

What we have not found is the retraction σ for an arbitrary finitely presented group: the blank third sector is what lets the inverse sector of V_s violate the source relators without being tested. It gives the upper bounds, and with them the comparison of the whole spectrum in both directions.

5.4 Linear system games

Three routes lead to linear system games. The first applies to virtual retracts of the presentations, with at most one non-involutive generator, for which our companion paper writes out Slofstra's embedding [101, Appendix B], and keeps the state in both directions; the second applies to Slofstra's extended presentations with any number of non-involutive generators and keeps it in the lower direction; the third, the restricted compiler behind the explicit exported games, is described in Section 11.2.

Virtual retracts of $E\Gamma_0$ presentations. A *group over* $\mathbb{Z}_2$ is presented as $\langle S : R\rangle$ with a distinguished central involution $J \in S$ and $J^2, [J,s] \in R$. Let $\mathcal{Q} = \langle S_0 \sqcup S_1 : R_{\mathcal{Q}}\rangle$ be an $E\Gamma_0$ presentation in the sense of [101, Definition 10.1], following [132, Definition 4.7], with group H, and $w \in S_0$ central in H. We call G a *generator-preserving virtual retract of* $(\mathcal{Q}, w)$ *of index* k if there are $H_0 \leq H$ of index k and homomorphisms $\iota : G \to H_0$ and $\pi : H_0 \to G$ with $\pi \circ \iota = \mathrm{id}$, such that each $\iota(s)$ is a generator $x_s \in S_0 \sqcup S_1$, with $x_s \neq x_{s'}$ for $s \neq s'$ and $x_J = w$. Suppose that $|S_1| \leq 1$ and that every linear set of $\mathcal{Q}$ has at most three elements; reordering the product of a linear set, whose elements commute, presents the same group, so we may also assume the ordering hypothesis of [101, Lemma B.1]. That lemma, the explicit form of Slofstra's embedding [132, Section 4], gives a binary linear system (A, b) with solution group Γ, a homomorphism Ψ from the free group on $S_0 \sqcup S_1$ to $F(x_1, \dots, x_n, J)$ and a variable x_{j_0} such that $\Psi(r)$ is trivial in Γ for every $r \in R_{\mathcal{Q}}$ and $[\Psi(w), x_{j_0}]J$ is trivial in Γ.

Theorem 5.8. *Let G be a generator-preserving virtual retract of index k of $(\mathcal{Q}, w)$, where $\mathcal{Q}$ is an $E\Gamma_0$ presentation with $|S_1| \leq 1$ and linear sets of at most three elements, and let Γ be the solution group of [101, Lemma B.1] for $\mathcal{Q}$ and w. Write h_G, h_G^S and h_Γ, h_Γ^S for the profiles of J. There are constants $C, C' > 0$ such that, for every $e > 0$,*

$$h_G(Ce) \leq h_\Gamma(e) \leq 8k\, h_G(e/C'), \qquad h_G^S(Ce) \leq h_\Gamma^S(e) \leq h_G^S(e/C') + \log(8k).$$

If $h_G(e) < \infty$ for every $e > 0$, the game $G(A, b)$ has quantum value one, and there are $C_1, C_2 > 0$ with

$$h_G(C_1\sqrt{\varepsilon}) \leq \mathcal{E}(G(A,b), \varepsilon) \leq 8k\, h_G(C_2\sqrt{\varepsilon})$$

for every $\varepsilon > 0$; the same holds for the least entanglement entropy of pure strategies, with h_G^S in place of h_G and $\log(8k)$ added on the right.

Proof. Upper bounds. Let $t_1 = e, \dots, t_k$ be coset representatives of H_0, and $c(h, i) = t_{h\cdot i}^{-1} h t_i \in H_0$ the induction cocycle. For each generator x of $\mathcal{Q}$ and each i fix a word $\beta(x, i)$ for $\pi(c(x, i))$, with $\beta(w, i) = J$; this is possible because $w = \iota(J)$ is central and lies in H_0, so that $w \cdot i = i$ and $c(w, i) = t_i^{-1} w t_i = w$, whose image $\pi(w) = \pi(\iota(J))$ is J. The monomial matrices $\sigma_0(x)$ with entry $\beta(x, i)$ in position $(x \cdot i, i)$ define a homomorphism from H into the monomial matrices over G, by the cocycle identity, and so a unital $*$-homomorphism $\sigma_0 : A_{\mathcal{Q}}^- \to M_k(A_G^-)$ with $\sigma_0(w) = -1$. The four steps of the proof of [101, Lemma B.1] pass from $\mathcal{Q}$ through presentations G', $\widehat{G}$ and $\widehat{G}_1$ to Γ. Each sends every generator of the next presentation to

a block-monomial matrix with entries $\pm$ words in the previous generators, or constants; the first, third and fourth steps double the number of blocks. The proof there shows that the image of every relator, apart from the relators Jwu and $t^*wt^* = u$ of Step 2, is block diagonal with block words that are trivial in the previous group [101, proof of Lemma B.1, and Table 3]. Step 1 is a unital $*$-homomorphism $\mathbb{C}[G'] \to M_2(\mathbb{C}[H])$, but not a map of negative quotients: it sends each $x \in S_0$ to $\mathrm{diag}(x, 1)$, so w goes to $\mathrm{diag}(-1, 1)$ in $M_2(A_{\mathcal{Q}}^-)$. Step 2 adds involutions t^*, J, u, sent to the matrix that exchanges the two blocks, to -1 and to $-\mathrm{diag}(w, 1)$; its relators Jwu and $t^*wt^* = u$ hold because $w = -1$ in $A_{\mathcal{Q}}^-$. So Steps 1 and 2 together give a unital $*$-homomorphism $A_{\widehat{G}}^- \to M_2(A_{\mathcal{Q}}^-)$ with $J \mapsto -1$. Steps 3 and 4 send J to $J \oplus J$ and give unital $*$-homomorphisms $A_{\widehat{G}_1}^- \to M_2(A_{\widehat{G}}^-)$ and $A_\Gamma^- \to M_2(A_{\widehat{G}_1}^-)$. Their composite with σ_0 is a unital $*$-homomorphism $A_\Gamma^- \to M_{8k}(A_G^-)$, given by finitely many words, and Proposition 5.3 with $N = 8k$ gives the upper bounds. Steps 3 and 4 are the lifts of [132, Lemma 4.4 and Proposition 4.2], which turn a d-dimensional approximate representation into a $4d$-dimensional one restricting to four copies of it [132, Remark 4.5]; see also [60, Lemma 5.8 and Theorem 5.5].

Lower bounds. Let $\hat{\iota}(s) = x_s$ and $\psi = \Psi \circ \hat{\iota}$. It induces a homomorphism $G \to \Gamma$ that sends J to $\Psi(w)$, and $t = x_{j_0}$ satisfies $[\Psi(w), t] = J$ in Γ. As J is central in G, Proposition 5.5 with $k = 1$ gives the lower bounds.

The game. If $h_G(e) < \infty$ for every e, so is $h_\Gamma(e)$, and Theorem 4.5 gives strategies whose losing probability tends to 0, so $\omega_q = 1$. The two displayed game bounds follow from Corollary 4.6 and the profile bounds, and likewise for entropy. □

For the flat profile and $k = 1$, the companion's embedding theorem gives the two directions with explicit constants [101, Theorem B.3]; Theorem 5.8 adds virtual retracts, keeps the state in both directions, includes the entropy, and works at the natural precision $\sqrt{\varepsilon}$.

Remark 5.9 (A complete instance). For the group generated by involutions a, b, c and an element y with the relations $abc = 1$, $[a, b] = [a, c] = [b, c] = 1$, $yay^{-1} = b$, $yby^{-1} = a$ and $ycy^{-1} = c$, and with distinguished involution c, which is a generator-preserving retract of itself, the construction of [101, Lemma B.1], with the commutations inside the linear set kept as relators of the solution group rather than repeated as conjugacy relations, gives a linear system with 223 variables and 180 equations, as the counts of that lemma predict with no conjugacy relations among the involutions, and an 8-block lift of every variable. The script `verify/check_compiler.py` builds it, checks all 942 square, product, commutator and witness relations of the lift exactly in the negative quotient of the source, and runs the strategy of Theorem 4.5 on perturbed models. The group has exact finite-dimensional models, so this is a check of the construction, not a game with growing entanglement.

Extended presentations. Slofstra's *extended homogeneous-linear-plus-conjugacy presentations* [134, Section 6] (following [132]) have generators $S = S_0 \sqcup S_1$ with $S_1 = \{y_1, \dots, y_l\}$ linearly ordered; the relations $x^2 = e$ for $x \in S_0$; for each set V of a collection of ordered subsets of S_0, the relations $\prod_{x \in V} x = e$ and $xy = yx$ for distinct $x, y \in V$; and otherwise only relations $xyx = z$ with $x, y, z \in S_0$, $xyx^{-1} = z$ with $x \in S_1$ and $y, z \in S_0$, and $y_i y_j y_i^{-1} = y_j^{L_{ij}}$ with $j < i$ and $L_{ij} > 0$. Let $w \in S_0$ represent a central involution of G that is non-trivial in approximate representations. For such a presentation Slofstra and Vidick state that there are a solution group Γ and constants with $\mathrm{hlp}_\Gamma(J; 2, \varepsilon) \le N\ \mathrm{hlp}_G(w; 2, C\varepsilon)$ [134, Proposition 6.2]. Their proof prints $N = 8$: it takes from [132, Proposition 4.8] a $2d$-dimensional lift of a d-dimensional representation with $w \mapsto (-I_d) \oplus I_d$, and then adds an involution t and the central J with $twt = wJ$, realized by the exchange of the two blocks. That proposition eliminates the non-involutive generators one at a time, and each step doubles the dimension and lifts every involutive generator x as $\varphi(x) \oplus I$. For $l = 1$ this is the lift used. For $l \ge 2$ the

marked involution becomes $W = (-I_d) \oplus I_{(2^l-1)d}$, and no unitary T satisfies $TWT^* = -W$, even approximately: $\operatorname{Tr}(TWT^* + W) = 2 \operatorname{Tr} W = 2(2^l - 2)d$, so $\|TWT^* + W\|_f \geq 2(1 - 2^{1-l})$ by the Cauchy–Schwarz inequality, whatever the precision of φ. This concerns only the factor N of the upper bound: the lower bound of [134, Proposition 6.2], and the non-triviality of J in approximate representations of Γ, which [134] take from [132, Lemma 5.2], do not use this lift. The lamp sources of Section 11 have $l = 2$ or $l = 2k$, and the group behind the Slofstra–Vidick game has $l = 2$ (Section 9.4). We therefore use a variant that changes the marked element before the sign extension, and prove what we use.

The balanced construction. Let $l \geq 1$ and $k = 2^{l-1}$. Eliminating $y_1, \dots, y_l$ in order as in [132, Proposition 4.8] adds involutions f_j, v_j with $y_j = f_j v_j$ and gives a homogeneous-linear-plus-conjugacy presentation G', all of whose generators are involutions, with a homomorphism $G \to G'$ that is the identity on S_0. For $0 \leq i < k$ let g_i be the product of the f_j over the binary digits 2^{j-1} of i, in decreasing order of j; put $\psi_i(s) = g_i s g_i^{-1}$ for $s \in S$, read in G', and $W_i = \psi_i(w)$. Add to G' the relations that $W_0, \dots, W_{k-1}$ commute with each other and with every $\psi_j(s)$, $s \in S$, $j < k$; an involution p with the row $pW_0 \cdots W_{k-1} = e$; the central J; and an involution u with the row $pu = J$ and the conjugacy $f_l p f_l = u$. Every $\psi_j(s)$ is a word in involutions, so with auxiliary involutions for the conjugates, as in [132, Proposition 4.8], all new relations are rows and involutive conjugacies. The result $\widehat{G}$ is a linear-plus-conjugacy group, and [132, Lemma 4.4 and Proposition 4.2] give a solution group $\Gamma = \Gamma(A, b)$ and a homomorphism $\widehat{G} \to \Gamma$ fixing J. In the lift below the W_i are -1 on disjoint blocks, and the new marker p has its two eigenvalues with equal multiplicities.

Proposition 5.10. *Let G and w be as above, with $l \geq 1$ non-involutive generators, let $\Gamma = \Gamma(A, b)$ be the solution group of the balanced construction, and put $N_G = 2^{l+2}$. There are constants $C_1, C_2 > 0$ such that, for the game $G(A, b)$ with either question distribution of Definition 4.1 and every $\varepsilon > 0$:*

(a) every strategy that wins with probability at least $1 - \varepsilon$ has Schmidt rank at least $h_G(C_1\sqrt{\varepsilon})$, and the nonzero spectrum of Bob's reduced density matrix lies in $\Lambda_G(C_1\sqrt{\varepsilon})$;

(b) if $\operatorname{hlp}_G(w; 2, e) < \infty$ for every $e > 0$, then $\omega_q(G(A, b)) = 1$ and $\mathcal{E}(G(A, b), \varepsilon) \leq N_G \operatorname{hlp}_G(w; 2, C_2\sqrt{\varepsilon})$.

Proof. (a) In Γ each ψ_i induces a homomorphism $G \to \Gamma$, the $W_i = \psi_i(w)$, $i < k$, commute with each other and with every $\psi_j(s)$, and $b = W_0 \cdots W_{k-1}$ equals p, so $t = f_l$ satisfies $tbt^{-1} = u = Jb$. Theorem 4.3 gives a model of Γ at precision $O(\sqrt{\varepsilon})$ on the Schmidt support, with Bob's reduced state, and Proposition 5.5 with these k maps ψ_i and this t turns it into a model of G with the same state.

(b) Let φ be a d-dimensional e-representation of G with $\varphi(w) = -I$, and let γ be the $2^l d$-dimensional $O(e)$-representation of G' that the l lifts of [132, proof of Proposition 4.8] give: at step j, with ρ the representation reached so far, every involutive generator x of the current presentation becomes $\rho(x) \oplus I$, f_j exchanges the two summands, $v_j = \left(\begin{smallmatrix} 0 & \rho(y_j)^* \\ \rho(y_j) & 0 \end{smallmatrix}\right)$, and every later y_i becomes $\rho(y_i) \oplus \rho(y_i)$. Number the 2^l blocks of size d so that step j adds the blocks $2^{j-1}, \dots, 2^j - 1$. By induction on the steps, $\gamma(f_j)$ exchanges the blocks i and $i + 2^{j-1}$ for $i < 2^{j-1}$ and fixes the others; every generator in S, read in G', acts block-diagonally; and $\gamma(w)$ is $-I$ on block 0 and I on the others. Hence $\gamma(g_i)$ carries block 0 to block i, every $\gamma(\psi_j(s))$ is block-diagonal, and $\gamma(W_i)$ is $-I$ on block i and I elsewhere. Put $\gamma(p) = \gamma(W_0) \cdots \gamma(W_{k-1})$, which is $-I$ on the first k blocks and I on the others, $\gamma(u) = -\gamma(p)$ and $\gamma(J) = -I$, and give the auxiliary involutions their defining values. The added relations then hold exactly; in particular $\gamma(f_l)$ exchanges the two halves, so $\gamma(f_l)\gamma(p)\gamma(f_l) = -\gamma(p) = \gamma(u)$. So γ is an $O(e)$-representation of $\widehat{G}$ with $J \mapsto -I$, and [132, Remark 4.5] turns it into a $2^{l+2}d$-dimensional $O(e)$-representation of Γ with $J \mapsto -I$. So Γ has flat models with $J \mapsto -I$ at every precision,

of dimension at most N_G $\mathrm{hlp}_G(w; 2, Ce)$, and Theorem 4.5 with the flat state turns them into strategies. □

For $l = 1$ the construction gives $N_G = 8$, the value printed in [134]. A fixed N_G adds $\log N_G = l + 2$ to the entropies of the upper strategies and leaves every leading coefficient and exponent unchanged. The restricted compiler of Section 11.2 has the same architecture with its own conversion stages; Proposition 5.10 does not certify its specifications. For the group of [134, Proposition 6.2] itself the lower direction needs only the relation $[\Psi(w), z] = J$ in its solution group, which holds there; Section 9 uses it with the corner map for $k = 1$. The upper bound for that group comes from a lift specific to its models, proved in our companion paper [101, Lemma 11.13] (Section 9.4).

6 Operational extraction on the original state

The counting theorem bounds the entropy of Bob's reduced state, which for a mixed strategy is not a bound on usable entanglement (Section 1.6). This section supplies the operational tools. Each takes approximate relations among actual operators of the two players, measured on a purification Ψ_{ABE} of the original mixed state, and produces local isometries whose output either is close to a maximally entangled state or has few Bell errors, so that hashing applies. Nothing acts on E, no pure-state ensemble decomposition is used, and the state is never replaced. The mechanisms (regular dilations, coherent Pauli extraction, pretty-good recovery, hashing) are established; what is new is the accounting that keeps a growing register on the original state with errors summed in squares [66, 107, 98, 51].

6.1 The actual source interface of a compiled game

A source enters a game through a compiler whose lower map sends each source generator s to a fixed polynomial F_s in the game's variables (Section 5). On a strategy Bob evaluates F_s on his variable reflections. Alice evaluates the *opposite* polynomial $\theta(F_s)$, obtained by reversing every word and substituting, for each variable j, her coarse-grained reflection for j in one fixed incident context $i(j)$. The map θ is a complex-linear $*$-preserving antihomomorphism of word polynomials, and Lemma 2.11(a) gives, for every polynomial $P = \sum_w c_w w$,

$$\|(P(B) - \theta(P)(A))\Psi\| \le \delta \sum_w |c_w||w|, \qquad \delta = \max_j \|(A_j - B_j)\Psi\|. \tag{6.1}$$

Proposition 6.1 (Actual source interface)**.** *Let G be a finite Boolean constraint-variable game, every incidence tested with probability at least $\pi_{\min} > 0$, and let $s \mapsto F_s$ be a lower map from a presentation $\langle S \mid R\rangle$ with marked central involution J into the game algebra $\mathcal{A}(G)$ of Example 5.1, given by fixed polynomials such that F is a unital $*$-homomorphism of the negative quotient. There is a constant C_G such that every finite-dimensional mixed strategy of loss ε, with purification Ψ after local projective dilation, has Bob unitaries U_s and Alice unitaries U_s^A with, for $e = C_G\sqrt{\varepsilon}$,*

(a) $\|(U_s^A - U_s)\Psi\| \le e$ *and* $\|(U_s^{A*} - U_s^*)\Psi\| \le e$ *for every s;*

(b) $\|(r(U) - I)\Psi\| \le e$ *for every relator r, and* $\|(r^{\mathrm{rev}}(U^A) - I)\Psi\| \le e$ *for the reversed word evaluated on Alice;*

(c) $\|(U_J + I)\Psi\| \le e$*;*

(d) $\|[U_s, \rho_B^{1/2}]\|_2 \le e$ *and* $\|[U_s^A, \rho_A^{1/2}]\|_2 \le e$*.*

If the F_s are words in the variables, one may take $U_s = F_s(B)$ and $U_s^A = \theta(F_s)(A)$ without rounding.

Proof. Relabel every illegal Alice answer to one fixed legal answer of its context, which cannot increase the loss, and dilate all measurements to projective ones with local pure ancillas, which changes neither the correlations nor the directional rates. For a context i Alice's coarse-grained reflections A_{ij} then commute exactly and every forbidden ordered product of their spectral projections vanishes. A tested incidence is lost with probability at most $\varepsilon/\pi_{\min}$, so $\|(A_{ij} - B_j)\Psi\| \le \delta := 2\sqrt{\varepsilon/\pi_{\min}}$. Every game relator Q (a context commutator or a forbidden ordered product) is a word of bounded length in the variables of one context i; transferring it to that context's own Alice reflections by Lemma 2.11(a), where it is zero or the identity exactly, gives $\|Q(B)\Psi\| = O(\delta)$.

Every polynomial in the ideal of the game relations has a finite certificate $\sum_l a_l Q_l b_l$. For a term aQb, move $b(B)$ to Alice by (6.1), commute it past $Q(B)$, and use $\|\theta(b)(A)\| \le \sum |c_w|$: $\|a(B)Q(B)b(B)\Psi\| = O(\delta)$. Hence every element of the ideal has Bob residual $O(\delta)$ on Ψ, with a constant depending only on the fixed certificate, and by (6.1) its opposite polynomial has Alice residual $O(\delta)$. In particular this holds for $F_s^*F_s - 1$, $F_sF_s^* - 1$, for $r(F) - 1$ with $r \in R$, and for $F_J + 1$. Since θ reverses products, $\theta(F_s^*F_s) = \theta(F_s)\theta(F_s)^*$ and the two unitarity defects exchange roles on Alice; both are controlled.

Let U_s be a unitary polar part of $F_s(B)$ and U_s^A one of $\theta(F_s)(A)$. The scalar inequality $|1-t| \le |1-t^2|$, $t \ge 0$, gives $\|(U-F)\Psi\| \le \|(F^*F-I)\Psi\|$ and $\|(U^*-F^*)\Psi\| \le \|(FF^*-I)\Psi\|$ on any vector (the argument of Lemma 2.7, which uses only functional calculus of $|F|$). So U_s and U_s^A are within $O(\delta)$ of $F_s(B)$ and $\theta(F_s)(A)$, together with their adjoints, and (a) follows from (6.1). For (b) and (c), replace the letters of $r(U)$ one at a time, from the right, by the polynomials $F_s(B)$ or $F_s(B)^*$. At each step the already replaced suffix is a polynomial $S(B)$ of fixed coefficient norm and degree, not necessarily unitary on the strategy. Move it to Alice by (6.1), at cost $O(\delta)$ because the replacement error to its left is bounded. Then $\theta(S)(A)$, of norm at most the coefficient norm of S, commutes with that replacement error, which is $O(\delta)$ on Ψ. So each step costs $O(\delta)$, and $r(F) - 1$ has residual $O(\delta)$. The reversed Alice words are handled in the same way, with the parties exchanged. Part (d) is Lemma 2.11(d). All constants depend only on G, the presentation and the chosen certificates. □

The same proof applies to linear system games, whose game relations are the parity rows and context commutators, and to the corner maps of Proposition 5.5. The orientation matters: Alice realizes the opposite algebra. If Bob has $TY = \zeta YT$, Alice has $T_AY_A = \zeta^{-1}Y_AT_A$ in the corresponding reversed words, and the invariant tensor of a generator is $U_s^{A*} \otimes U_s$, not $U_s^A \otimes U_s$.

Selector polynomials. The proof never uses that F_s is the image of a generator, only that its unitarity defects lie in the game ideal. Let T be a fixed polynomial in the source generators such that $T^*T - 1$ and $TT^* - 1$ lie in the defining ideal of the negative quotient, and let T_F be the polynomial in the game variables obtained by substituting F_s for each generator s. Since F is a homomorphism of the negative quotient, the two unitarity defects of T_F lie in the game ideal. The certificate and polar argument above then gives, verbatim: Bob and Alice residuals $O(\delta)$ for the unitarity defects of T_F and of $\theta(T_F)$; the mirror $\|(T_F(B) - \theta(T_F)(A))\Psi\| = O(\delta)$, by (6.1); unitary polar parts U_T of $T_F(B)$ and U_T^A of $\theta(T_F)(A)$, within $O(\delta)$ of them together with their adjoints, so that U_T^A is an actual mirror of U_T, and U_T^{A*} of U_T^*, with error $O(e)$; square-root-state commutators $O(e)$, by Lemma 2.11(d); and residual $O(e)$ for every word in U_T, U_T^* and the U_s, U_s^* whose image in the polynomials lies in the game ideal, replacing letters from the right as in (b), with the reversed words on Alice. In particular, if $Ty = \zeta yT$ in the negative quotient for a generator y, then $\|(U_TU_yU_T^* - \zeta U_y)\Psi\| = O(e)$, and on Alice $\|(U_T^AU_y^AU_T^{A*} - \zeta^{-1}U_y^A)\Psi\| = O(e)$; and if T commutes with a generator g in the negative quotient, then $\|[U_T, U_g]\Psi\|$ and $\|[U_T^A, U_g^A]\Psi\|$ are $O(e)$. The constants depend only on G, the

presentation, T and the chosen certificates. This applies to the corner maps and linear system games as well, and it is used for the rotation selectors of Lemma 10.3 in Theorem 11.9.

6.2 Triangular extraction

Theorem 6.2 (Triangular extraction)**.** *Let Ψ be a unit vector in $H_A \otimes H_B \otimes H_E$ and $M \geq 1$. On each party let Z_j, T_j, $0 \leq j < M$, be reflections such that, for some $\delta \geq 0$,*

$$\|(Z_j^A - Z_j^B)\Psi\| \leq \delta, \qquad \|(T_j^A - T_j^B)\Psi\| \leq \delta,$$
$$\|[Z_j^A, Z_s^A]\Psi\| \leq \delta, \quad \|[T_j^A, Z_s^A]\Psi\| \leq \delta \quad (s < j), \qquad \|\{T_j^A, Z_j^A\}\Psi\| \leq \delta.$$

No relation between different T_j is assumed. There are product local isometries W_A, W_B and a unit vector χ, including E, with

$$\|(W_A \otimes W_B \otimes I_E)\Psi - \Phi_{2^M} \otimes \chi\| \leq \min\{2, 20M^3\delta\}.$$

Proof. Histories. Put $P_j^a = \frac{1}{2}(I + (-1)^a Z_j)$ on each party. For $h = (a_0, \ldots, a_{j-1})$ let $K_h^A = P_{j-1}^{A,a_{j-1}} \cdots P_0^{A,a_0}$ (chronological order) and $Q_h^B = P_0^{B,a_0} \cdots P_{j-1}^{B,a_{j-1}}$ (the reversed order). Summing the innermost equal pair of projections repeatedly gives $\sum_h K_h^* K_h = \sum_h K_h K_h^* = I$, and the same for Q, with no commutation. Put $\Delta_j = (\sum_h \|(K_h^A - Q_h^B)\Psi\|^2)^{1/2}$ and $z_s = \|(Z_s^A - Z_s^B)\Psi\|$. From $P_s^A K_h^A - Q_h^B P_s^B = P_s^A(K_h^A - Q_h^B) + Q_h^B(P_s^A - P_s^B)$, summing the new outcome with $\sum_a (P_s^a)^2 = I$ in the first term and the old histories with $\sum_h Q_h^* Q_h = I$ in the second, Minkowski's inequality gives $\Delta_{s+1} \leq \Delta_s + z_s/\sqrt{2}$, hence

$$\Delta_j \leq \sum_{s<j} z_s/\sqrt{2} \leq j\delta/\sqrt{2}. \tag{6.2}$$

For an operator F on Alice, $FK_h^A\Psi = Q_h^B F\Psi + F(K_h^A - Q_h^B)\Psi$, so

$$\Big(\sum_h \|FK_h^A\Psi\|^2\Big)^{1/2} \leq \|F\Psi\| + \|F\|_{\mathrm{op}}\Delta_j, \tag{6.3}$$

and for a finite family F_a the same holds with $(\sum_a \|F_a\Psi\|^2)^{1/2} + (\sum_a \|F_a\|_{\mathrm{op}}^2)^{1/2}\Delta_j$. There is no factor of the number of histories.

Commutators along histories. For a reflection G on Alice put $C_{G,j} = (\sum_h \|[G, K_h^A]\Psi\|^2)^{1/2}$. Telescoping $[G, P_{j-1} \cdots P_0]$, summing the later outcomes by completeness and applying (6.3) to the pair $F_a = [G, P_s^a] = \frac{(-1)^a}{2}[G, Z_s]$, of operator norm at most one,

$$C_{G,j} \leq \sum_{s<j} (\|[G, Z_s^A]\Psi\|/\sqrt{2} + \sqrt{2}\Delta_s) \leq j\delta/\sqrt{2} + j(j-1)\delta/2 \leq (j+1)^2\delta - \delta$$

for $G = Z_j$ or T_j. The mirror defect $G_A - G_B$ is bipartite and is not substituted into (6.3); instead $(G_A - G_B)K_h^A\Psi = K_h^A(G_A - G_B)\Psi + [G_A, K_h^A]\Psi$, so

$$\Big(\sum_h \|(G_A - G_B)K_h^A\Psi\|^2\Big)^{1/2} \leq (j+1)^2\delta \qquad (G = Z_j, T_j). \tag{6.4}$$

The conjugacy defect $C_j = T_j Z_j T_j + Z_j = T_j\{T_j, Z_j\}$ on Alice has norm at most two and $\|C_j\Psi\| \leq \delta$, so by (6.3) its aggregate over histories is at most $\delta + 2\Delta_j \leq (2j+1)\delta$.

One stage. For a target bit a let $W_j^a = \sum_{b=0,1} |b\rangle \otimes T_j^{a\oplus b} P_j^b$ on each party; it is an isometry because T_j is unitary. Consider

$$\mathcal{A}_j\eta = \tfrac{1}{\sqrt{2}} \sum_a |a, a, a\rangle \otimes (W_j^{A,a} \otimes W_j^{B,a})\eta, \qquad \mathcal{I}_j\eta = \sum_a |a, a, a\rangle \otimes \Phi_2 \otimes P_j^{A,a}\eta.$$

The first is the Stinespring form of a shared uniform random target, one copy of which is kept by the environment, followed by local isometries; the second is an exact isometry used only in the analysis. For any vector η put $t = \|(T_j^A - T_j^B)\eta\|$, $z = \|(Z_j^A - Z_j^B)\eta\|$ and $c = \|C_j\eta\|$. The branch of $\mathcal{A}_j\eta$ with flags (a, b, c') is $2^{-1/2}T_A^{a\oplus b}P_A^b T_B^{a\oplus c'}P_B^{c'}\eta$; transferring the Bob factor to Alice by Lemma 2.11(a) costs $t + z/2$ and gives $2^{-1/2}T^{a\oplus b}P^bP^{c'}T^{a\oplus c'}\eta$ on Alice. If $b \neq c'$ this is exactly zero, as $P^bP^{c'} = 0$; if $b = c' = a$ it is $P^a\eta$; and if $b = c' \neq a$ it is $TP^{1-a}T\eta = P^a\eta \mp \frac{1}{2}C_j\eta$. The ideal branch is $2^{-1/2}1_{b=c'}P_A^a\eta$. Over the eight orthogonal branches,

$$\|(\mathcal{A}_j - \mathcal{I}_j)\eta\| \leq 2t + z + c. \tag{6.5}$$

Telescoping. Compose the stages $j = 0, \ldots, M-1$ and compare $\mathcal{A}_{M-1}\cdots\mathcal{A}_0$ with $\mathcal{I}_{M-1}\cdots\mathcal{I}_0$ by telescoping, with actual maps on the left (isometries) and ideal prefixes on the right. The ideal prefix after j stages is a direct sum over histories h of $\Phi_2^{\otimes j} \otimes K_h^A\Psi$, so by (6.5), (6.4) and the conjugacy aggregate, the j-th term is at most $3(j+1)^2\delta + (2j+1)\delta \leq 5(j+1)^2\delta$. Summing, the actual composite is within $5M^3\delta$ of $\Phi_{2^M} \otimes \chi$, $\chi = \sum_a |a, a, a\rangle \otimes K_a^A\Psi$, a unit vector by completeness.

A deterministic target. For each target string a let v_a be the output of the local isometries with that target, a unit vector. The squared error just bounded is $2^{-M}\sum_a \|v_a - \Phi_{2^M} \otimes 2^{M/2}K_a^A\Psi\|^2$, so some a has $\|v_a - \Phi_{2^M} \otimes w\| \leq 5M^3\delta$ with $w = 2^{M/2}K_a^A\Psi$; if this is below one, replacing w by $w/\|w\|$ at most doubles it. The trivial bound 2 covers the remaining case. □

The triangular matrix-unit identities hold exactly in the source algebra but only approximately on the actual state, and the reversed orientation of the histories on the two parties is essential; with the same orientation on both sides the estimate (6.2) fails.

The same argument works with cyclic labels of bounded size and unitary controls of arbitrary order. Errors indexed by labels are measured by their *root aggregates*, the square roots of the sums of their squares over the labels.

Lemma 6.3 (Cyclic triangular extraction)**.** *Let Ψ be a unit vector in $H_A \otimes H_B \otimes H_E$, $Q \geq 2$ and $M \geq 1$. For $0 \leq j < M$ let $2 \leq q_j \leq Q$, and on each party $P \in \{A, B\}$ let $\{P_{j,a}^P : a \in \mathbb{Z}_{q_j}\}$ be a projective partition and U_j^P a unitary. Suppose that, for some $\delta \geq 0$, all of the following root aggregates on Ψ are at most δ, for every j, every $t < j$ and every integer h with $|h| \leq Q$:*

(i) the mirrors $(\sum_a \|(P_{j,a}^A - P_{j,a}^B)\Psi\|^2)^{1/2}$ and $\|((U_j^A)^h - (U_j^B)^h)\Psi\|$;

(ii) on Alice, $(\sum_{a,b} \|[P_{j,a}^A, P_{t,b}^A]\Psi\|^2)^{1/2}$ and $(\sum_b \|[(U_j^A)^h, P_{t,b}^A]\Psi\|^2)^{1/2}$;

(iii) on Alice, $(\sum_a \|((U_j^A)^h P_{j,a}^A (U_j^A)^{-h} - P_{j,a+h}^A)\Psi\|^2)^{1/2}$, the labels taken modulo q_j.

No relation between controls of different stages and no relation $(U_j^P)^{q_j} = I$ is assumed. There are a constant C_Q, depending only on Q, product local isometries W_A, W_B and a unit vector χ, including E, with

$$\|(W_A \otimes W_B \otimes I_E)\Psi - \Phi_{q_0q_1\cdots q_{M-1}} \otimes \chi\| \leq \min\{2, C_Q M^3\delta\}.$$

Proof. We follow the proof of Theorem 6.2 and indicate the changes; C_Q denotes constants depending only on Q.

Histories. For a history $h = (a_0, \ldots, a_{j-1})$ put $K_h^A = P_{j-1,a_{j-1}}^A \cdots P_{0,a_0}^A$ and $Q_h^B = P_{0,a_0}^B \cdots P_{j-1,a_{j-1}}^B$. Summing the innermost partition repeatedly gives $\sum_h K_h^*K_h = \sum_h K_hK_h^* = I$, and the same for Q, with no commutation. With $\Delta_j = (\sum_h \|(K_h^A - Q_h^B)\Psi\|^2)^{1/2}$, the identity $P_{j,a}^A K_h^A - Q_h^B P_{j,a}^B = P_{j,a}^A(K_h^A - Q_h^B) + Q_h^B(P_{j,a}^A - P_{j,a}^B)$, completeness

in each term and Minkowski's inequality give $\Delta_{j+1} \le \Delta_j + \delta$, so $\Delta_j \le j\delta$. As in (6.3), for a finite family F_b of Alice operators,

$$\Big(\sum_{h,b} \|F_b K_h^A \Psi\|^2\Big)^{1/2} \le \Big(\sum_b \|F_b \Psi\|^2\Big)^{1/2} + \Big(\sum_b \|F_b\|_{\mathrm{op}}^2\Big)^{1/2} \Delta_j. \tag{6.6}$$

Commutators and mirrors along histories. Let G be $(U_j^A)^n$ with $|n| \le Q$, or the family $(P_{j,a}^A)_a$. Telescoping $[G, K_h^A]$ over the earlier stages $s < j$, summing the later outcomes by completeness, and applying (6.6) to the family $[G, P_{s,b}^A]$, whose operator norms are at most two and whose number is at most Q^2, gives by (ii)

$$\Big(\sum_h \|[G, K_h^A]\Psi\|^2\Big)^{1/2} \le \sum_{s<j} (\delta + 2Q\Delta_s) \le C_Q (j+1)^2 \delta,$$

summed over a as well when G is the family. Since the Bob operator G_B commutes with K_h^A, $(G_A - G_B)K_h^A\Psi = K_h^A(G_A - G_B)\Psi + [G_A, K_h^A]\Psi$, and completeness and (i) give the history aggregate $C_Q(j+1)^2\delta$ for the mirrors of $P_{j,\cdot}$ and of $(U_j)^n$. The covariance defects of (iii) have operator norm at most two, so by (6.6) their history aggregate is at most $C_Q(j+1)\delta$.

One stage. Drop the index j. For a target label $t \in \mathbb{Z}_q$ put

$$W_A^t = \sum_{b \in \mathbb{Z}_q} |b\rangle \otimes (U^A)^{t-b} P_b^A, \qquad W_B^t = \sum_{c \in \mathbb{Z}_q} |c\rangle \otimes (U^B)^{c-t} P_c^B,$$

with exponents taken in $\{-q+1, \dots, q-1\}$; each is an exact isometry, since the recorded label makes the cross terms orthogonal. Consider

$$\mathcal{A}\eta = q^{-1/2} \sum_t |t,t,t\rangle \otimes (W_A^t \otimes W_B^t)\eta, \qquad \mathcal{I}\eta = \sum_t |t,t,t\rangle \otimes \Phi_q \otimes P_t^A \eta;$$

the first is the Stinespring form of a shared uniform random target, one copy of which is kept by the environment, followed by local isometries, and the second is an exact isometry used only in the analysis. The branch of $\mathcal{A}\eta$ with flags (t,b,c) is $q^{-1/2}(U^A)^{t-b}P_b^A(U^B)^{c-t}P_c^B\eta$, and the ideal branch is $q^{-1/2}1_{b=c}P_t^A\eta$. Transferring the Bob word to Alice in reverse order,

$$(U^B)^{c-t}P_c^B\eta = P_c^A(U^A)^{c-t}\eta + (U^B)^{c-t}(P_c^B - P_c^A)\eta + P_c^A((U^B)^{c-t} - (U^A)^{c-t})\eta,$$

splits each branch, after the Alice factor $q^{-1/2}(U^A)^{t-b}P_b^A$, into three parts. The first is $q^{-1/2}(U^A)^{t-b}P_b^AP_c^A(U^A)^{c-t}\eta$. It vanishes exactly for $b \ne c$, and for $b = c$ it is $q^{-1/2}(U^A)^{t-b}P_b^A(U^A)^{b-t}\eta$, which differs from the ideal branch by $q^{-1/2}$ times a covariance defect of (iii) with exponent $n = t - b$ and label b. The second, the projection-mirror part, is $q^{-1/2}(U^A)^{t-b}P_b^A v_{tc}$ with $v_{tc} = (U^B)^{c-t}(P_c^B - P_c^A)\eta$. The unrecorded vectors $(U^A)^{t-b}P_b^A v$ need not be orthogonal as b varies (for $q = 2$, $t = 0$, $U^A = X$, $P_b^A = |b\rangle\langle b|$ and $v = |+\rangle$ both are $|0\rangle/\sqrt{2}$), but unitarity and completeness give

$$\sum_b \|(U^A)^{t-b}P_b^A v\|^2 = \sum_b \|P_b^A v\|^2 = \|v\|^2$$

for every vector v. So the projection-mirror parts have squared norms summing over (t,b,c) to $\sum_c \|(P_c^A - P_c^B)\eta\|^2$. The third, the unitary-mirror part, vanishes unless $b = c$, and the pairs (t,c) with $c - t = n$ number $q - |n| \le q$, so these parts have squared norms summing to at most $\sum_{|n|<q} \|((U^A)^n - (U^B)^n)\eta\|^2$. In the same way the covariance parts have squared norms summing to at most $q^{-1}\sum_{|n|<q}\sum_b \|((U^A)^n P_b^A (U^A)^{-n} - P_{b+n}^A)\eta\|^2$. The output branches carry orthogonal recorded labels (t,b,c), so $\|(\mathcal{A} - \mathcal{I})\eta\|^2$ is the sum of the squared branch

errors, and Minkowski's inequality for the projection-mirror, unitary-mirror and covariance contributions gives

$$\|(\mathcal{A}-\mathcal{I})\eta\| \le C_Q\Big[\Big(\sum_c \|(P_c^A - P_c^B)\eta\|^2\Big)^{1/2} + \sum_{|n|<q} \|((U^A)^n - (U^B)^n)\eta\| + \sum_{|n|<q}\Big(\sum_b \|((U^A)^n P_b^A (U^A)^{-n} - P_{b+n}^A)\eta\|^2\Big)^{1/2}\Big], \quad (6.7)$$

here with $C_Q = 1$.

Telescoping and a deterministic target. The ideal prefix after j stages is a direct sum over histories h of $\Phi_{q_0\cdots q_{j-1}} \otimes K_h^A\Psi$, with orthogonal flags, so the squared error of stage j on it is the sum over h of the squared errors (6.7) at $\eta = K_h^A\Psi$. Minkowski's inequality over the at most $4Q$ terms of (6.7), with the history aggregates, bounds the j-th term of the telescoping sum by $C_Q(j+1)^2\delta$, so the actual composite is within $C_Q M^3\delta$ of $\Phi_{q_0\cdots q_{M-1}} \otimes \chi$, $\chi = \sum_h |h,h,h\rangle \otimes K_h^A\Psi$, a unit vector by completeness; here $|h,h,h\rangle$ records the target string in three copies. For each target string t let v_t be the output of the local isometries with that target. The squared error just bounded is $(q_0\cdots q_{M-1})^{-1}\sum_t \|v_t - \Phi\otimes(q_0\cdots q_{M-1})^{1/2}K_t^A\Psi\|^2$, so some t has $\|v_t - \Phi\otimes w\| \le C_Q M^3\delta$ with $w = (q_0\cdots q_{M-1})^{1/2}K_t^A\Psi$; if this is below one, normalizing w at most doubles it, and the trivial bound 2 covers the remaining case. □

For $q_j = 2$ and reflections U_j the hypotheses agree with those of Theorem 6.2 up to constant factors, with $Z_j = P_{j,0} - P_{j,1}$: the covariance (iii) has aggregate $\|\{U_j, Z_j\}\Psi\|/\sqrt{2}$. The lemma is the finite-alphabet form of that theorem, not a new self-testing or hashing mechanism.

6.3 Signed regular extraction

Let $Q_1,\dots,Q_{2N}$ be the labels $X_1,\dots,X_N,Z_1,\dots,Z_N$, and let $s_{jk} = -1$ if $\{Q_j,Q_k\}$ is a pair $\{X_i,Z_i\}$ and $s_{jk} = +1$ otherwise. On the label register $\ell^2(\{0,1\}^{2N})$ define

$$R_k|v\rangle = r_k(v)|v+e_k\rangle, \qquad r_k(v) = \prod_{j>k} s_{jk}^{v_j}. \quad (6.8)$$

These are real reflections with $R_jR_k = s_{jk}R_kR_j$ exactly. Writing $v = (a,b)$, $R_{X_i} = S_{a_i}F_{b_i}$ and $R_{Z_i} = S_{b_i}$, with S a bit flip and F its sign; for one pair the real unitary $|z\rangle|t\rangle \mapsto |t+z\rangle H|z\rangle$ conjugates (R_X,R_Z) to $(X\otimes I, Z\otimes I)$. So the label algebra is exactly $M_2^{\otimes N}$ with multiplicity 2^N, by a fixed real identification.

Theorem 6.4 (Signed regular extraction)**.** *Let Ψ purify ρ_{AB}, and on each party $P\in\{A,B\}$ let $Q_1^P,\dots,Q_{2N}^P$ be reflections. Put*

$$\eta_{jk,P} = \|(Q_j^PQ_k^P - s_{jk}Q_k^PQ_j^P)\Psi\|, \quad \mu_j = \|(Q_j^A - Q_j^B)\Psi\|, \quad c_{j,P} = \|[Q_j^P, \rho_P^{1/2}]\|_2,$$

and $D_{k,P} = 4\sum_{j>k}[(\eta_{jk,P}+c_{j,P})^2 + c_{j,P}^2]$. There are exact local isometries V_A,V_B, with N output qubits per party after the fixed identification, whose k-th Bell stabilizer fails with probability

$$p_k \le \big(\sqrt{D_{k,A}} + \sqrt{D_{k,B}} + \mu_k\big)^2/4.$$

Consequently $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge N - \sum_{k=1}^{2N} h(p_k)$, where upper bounds may be substituted for the p_k in the monotone range $p_k \le \frac12$. If every signed pair error, state commutator and mirror error is at most κ, then $p_k \le 25N\kappa^2$, and

$$E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge N\big[1 - 2h(25N\kappa^2)\big] \qquad \text{when } 25N\kappa^2 \le \tfrac12.$$

No independence of the errors is assumed, and these are bounds for the original mixed state.

Proof. For $v \in \{0,1\}^{2N}$ put $f_P(v) = Q_1^{v_1} \cdots Q_{2N}^{v_{2N}}$ and $V_P u = 2^{-N} \sum_v f_P(v) u \otimes |v\rangle$, an exact isometry since each $f_P(v)$ is unitary. With $\xi = \rho_P^{1/2}$ and $d_{k,P} = \|(R_k V_P - V_P Q_k)\xi\|_2$, the coefficient of $|v\rangle$ is $2^{-N}(r_k(v) f(v+e_k) - f(v) Q_k)\xi$, and writing $f(v) = L Q_k^{v_k} T$ with the suffix $T = Q_{k+1}^{v_{k+1}} \cdots Q_{2N}^{v_{2N}}$, left unitary cancellation gives

$$d_{k,P}^2 = \mathbb{E}_v \|(r_k(v) T^* Q_k T - Q_k)\xi\|_2^2.$$

Put $u = Q_k \xi$. Then $(r_k T^* Q_k T - Q_k)\xi = (r_k T^* u T - u) + r_k T^* Q_k T(\xi - T^* \xi T)$. The signed conjugations $H \mapsto s_{jk} Q_j H Q_j$ are self-adjoint unitaries of the unweighted Hilbert–Schmidt space, so Lemma 2.9, with signs s_{jk}, gives $\mathbb{E}_v \|r_k T^* u T - u\|_2^2 \le \sum_{j>k} \|Q_j u - s_{jk} u Q_j\|_2^2$ and $\mathbb{E}_v \|T^* \xi T - \xi\|_2^2 \le \sum_{j>k} c_{j,P}^2$. Since $Q_j Q_k \xi - s_{jk} Q_k \xi Q_j = (Q_j Q_k - s_{jk} Q_k Q_j)\xi + s_{jk} Q_k [Q_j, \xi]$ and $\|(\cdot)\xi\|_2 = \|(\cdot)\Psi\|$ by Lemma 2.1,

$$d_{k,P}^2 \le 2 \sum_{j>k} (\eta_{jk,P} + c_{j,P})^2 + 2 \sum_{j>k} c_{j,P}^2 = \tfrac{1}{2} D_{k,P}. \tag{6.9}$$

For $\omega = (V_A \otimes V_B \otimes I_E)\Psi$, opposite-party maps commute, so $\|(R_k^A - R_k^B)\omega\| \le d_{k,A} + d_{k,B} + \mu_k$, and for opposite-party reflections the squared difference norm is four times the failure probability of their product. This proves the bound on p_k, a fortiori with $D_{k,P}$ in place of $D_{k,P}/2$. After the fixed real identification and discarding the local junk (the original systems and the label multiplicity), apply (2.3). In the uniform case use the sharper intertwiner estimate (6.9), $d_{k,P}^2 \le 2T_{k,P}$ with $T_{k,P} = \sum_{j>k} [(\eta_{jk,P} + c_{j,P})^2 + c_{j,P}^2]$, not the looser $D_{k,P} = 4T_{k,P}$ of the statement: there are fewer than $2N$ indices $j > k$, each with $(\eta_{jk,P} + c_{j,P})^2 \le (2\kappa)^2$ and $c_{j,P}^2 \le \kappa^2$, so

$$d_{k,P}^2 \le 2(2N)\big[(2\kappa)^2 + \kappa^2\big] = 20N\kappa^2, \qquad p_k \le \frac{(2\sqrt{20N} + 1)^2 \kappa^2}{4} \le 25N\kappa^2,$$

since $(2\sqrt{20N} + 1)^2 \le 100N$ for $N \ge 1$. The constant 25 is not obtained from $D_{k,P}$, which is already loosened by a factor two. □

The square-root-state commutators cannot be omitted; for reflections they are supplied by actual mirrors through Lemma 2.11(d), $c_{j,P} \le \sqrt{2}\mu_j$. A whole-register fidelity bound would add the $2N$ failure probabilities and lose a factor N; hashing uses only their entropy. For the qualitative law of Section 7 a fixed-size version suffices, which needs no state commutators.

Lemma 6.5 (Fixed-size signed regular extraction)**.** *Let $m = 2r$ and $C_r = m(1 + 2m)$. If reflections $Q_1^P, \dots, Q_m^P$ on each party satisfy $\eta_{jk,P} \le \kappa$ and $\mu_j \le \kappa$, then the isometries of Theorem 6.4 give Bell-stabilizer failures at most $p = (2C_r + 1)^2 \kappa^2 / 4$, singlet infidelity at most $2rp$, and a unit vector χ with $\|(V_A \otimes V_B \otimes I)\Psi - \Phi_{2^r} \otimes \chi\| \le \sqrt{4rp}$. If $p \le \frac{1}{2}$, then $E_{\mathrm{D}}^{A \to B}(\rho_{AB}) \ge r[1 - 2h(p)]$.*

Proof. Move Q_k through the suffix T letter by letter from the right. A swap with Q_j leaves the pair defect $Q_j Q_k - s_{jk} Q_k Q_j$, of norm at most two, followed by a suffix of at most m reflections; by Lemma 2.11(b) its contribution on Ψ is at most $\kappa + 2m\kappa$. So each label block of $(R_k V_P - V_P Q_k)\Psi$ has norm at most $C_r \kappa$ times its weight, and $\|(R_k V_P - V_P Q_k)\Psi\| \le C_r \kappa$. Hence $\|(R_k^A - R_k^B)\omega\| \le (2C_r + 1)\kappa$. The $2r$ Bell stabilizers commute and their common $+1$ space is Φ_{2^r}, so the infidelity is at most $2rp$, and the normalized projection of ω onto the Bell factor is within $\sqrt{2 \cdot 2rp}$ of ω. The last claim is (2.3). □

6.4 Aggregate and nonabelian bounds

Theorem 6.6 (Aggregate signed extraction)**.** *In the setting of Theorem 6.4 put* $E_B = \sum_{k<j} \eta_{jk,B}^2$ *and* $M_2 = \sum_j \mu_j^2$*. The average Bell-stabilizer failure of the extracted* N *pairs is at most*

$$b_N = 9E_B/N + 132M_2,$$

and if $b_N \le \frac{1}{2}$ *then* $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge N[1 - 2h(b_N)]$*. A negative lower bound may be replaced by zero. This is an entropy-density statement, not a full-register fidelity theorem at the same scale.*

Proof. Write $m = 2N$, $E_P = \sum_{k<j} \eta_{jk,P}^2$ and $C_P = \sum_j (j-1)c_{j,P}^2$. From the definition of $D_{k,P}$ and $(a+b)^2 \le 2a^2 + 2b^2$, $\sum_k D_{k,P} \le 8E_P + 12C_P$, and so $\sum_k p_k \le 6(E_A + E_B) + 9(C_A + C_B) + \frac{3}{4}M_2$. By Lemma 2.11(d), $c_{j,P} \le \sqrt{2}\mu_j$, hence $C_A + C_B \le 4(m-1)M_2$; transferring the two-letter products to the other party, $\eta_{jk,A} \le \eta_{jk,B} + 2(\mu_j + \mu_k)$, hence $E_A \le 2E_B + 16(m-1)M_2$. So $\sum_k p_k \le 18E_B + 132mM_2$. Divide by m, use concavity of h, and apply (2.3). □

Now let G be a finite group and, on each party, let $u_1^P, \dots, u_N^P$ be *exact* unitary representations of G, not assumed to commute for different i, oriented so that actual consistency reads

$$\|(u_i^A(g) - u_i^B(g^{-1}))\Psi\| \le \mu \qquad (g \in G), \tag{6.10}$$

which is the right orientation for nonabelian G, Alice representing the opposite algebra. On $\ell^2(G)$ use the right regular representation and a unitary Fourier decomposition $\ell^2(G) = \bigoplus_\lambda V_\lambda \otimes \mathbb{C}^{d_\lambda}$. For $d \ge 2$ let $F_d(t) = h(\min(t, 1-d^{-2})) + \min(t, 1-d^{-2})\log(d^2-1)$, the largest entropy of a d^2-valued Bell error of failure probability at most t, which saturates at $2\log d$; put $F_1 = 0$.

Theorem 6.7 (Finite-group extraction)**.** *Suppose, on both parties,* $\|[u_i^P(h), u_j^P(g)]\rho_P^{1/2}\|_2 \le \kappa$ *for* $i \ne j$, $\|[u_j^P(g), \rho_P^{1/2}]\|_2 \le \kappa$*, and* $\mu \le \kappa$*. There are exact local regular isometries and a one-way protocol in which Alice measures and announces the central types* α_i *of her label factors, such that, with* $a_{i\lambda}$ *the probability of type* λ *at coordinate* i *and* $q_{i\lambda} \le a_{i\lambda}$ *the unnormalized failure of the invariant projector there,*

$$E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \;\ge\; \sum_{i,\lambda} a_{i\lambda}[\log d_\lambda - F_{d_\lambda}(q_{i\lambda}/a_{i\lambda})], \qquad \sum_\lambda q_{i\lambda} \le 50N\kappa^2.$$

Zero-weight summands are zero, and Alice may discard any coordinate-type branch whose summand is negative. If only Bob cross-commutators are bounded, by η*, then* $\kappa = \eta + 4\mu$ *is admissible.*

Proof. For $x \in G^N$ put $f_P(x) = u_1^P(x_1)\cdots u_N^P(x_N)$ and $W_P v = |G|^{-N/2}\sum_x f_P(x) v \otimes |x\rangle$, an exact isometry. With $R_i(h)|x\rangle = |x_1, \dots, x_i h^{-1}, \dots, x_N\rangle$, the coefficient of $|x\rangle$ in $R_i(h)W_P$ is $f_P(xh_i)$, and for the suffix $T = u_{i+1}(x_{i+1})\cdots u_N(x_N)$ left cancellation gives $\delta_{i,P}(h)^2 := \|(R_i(h)W_P - W_P u_i(h))\xi\|_2^2 = \mathbb{E}_T\|(T^* u_i(h) T - u_i(h))\xi\|_2^2$. The maps $A_j(H) = \mathbb{E}_g u_j(g)^* H u_j(g)$ are orthogonal projections on the Hilbert–Schmidt space, and $\|(I - A_j)z\|^2 = \frac{1}{2}\mathbb{E}_g\|[u_j(g), z]\|_2^2$. The projection-product inequality (Lemma 2.8) gives $\mathbb{E}_T\|T^* z T - z\|_2^2 \le 2\sum_{j>i}\mathbb{E}_g\|[u_j(g), z]\|_2^2$. Applied to $z = u_i(h)\xi$ and to ξ, as in Theorem 6.4, this gives $\delta_{i,P}(h)^2 \le 20N\kappa^2$. For $\omega = (W_A \otimes W_B \otimes I)\Psi$, $\|(R_i^A(g) - R_i^B(g^{-1}))\omega\| \le (2\sqrt{20N} + 1)\kappa$. The diagonal invariant projector $\Pi_i = \mathbb{E}_g R_i^A(g) \otimes R_i^B(g)$ has failure $q_i = \langle\omega, (I - \Pi_i)\omega\rangle = \frac{1}{2}\mathbb{E}_g\|(R_i^A(g) - R_i^B(g^{-1}))\omega\|^2 \le 50N\kappa^2$.

Let $P_{i\lambda}^A$ be Alice's central type projections, $a_{i\lambda} = \langle\omega, P_{i\lambda}^A\omega\rangle$ and $q_{i\lambda} = \langle\omega, P_{i\lambda}^A(I - \Pi_i)\omega\rangle$; they commute, so $0 \le q_{i\lambda} \le a_{i\lambda}$ and $\sum_\lambda q_{i\lambda} = q_i$. By Schur's lemma, on the matching blocks $(\lambda, \bar\lambda)$

$$\Pi_i = \bigoplus_\lambda |\Phi_{d_\lambda}\rangle\langle\Phi_{d_\lambda}| \otimes I_{\text{multiplicities}},$$

and Π_i vanishes on the other pairs of types; Bob's matching type is the contragredient $\bar{\lambda}$. The protocol: both parties measure their central types; Alice announces her string α; Alice keeps the payload V_{α_i} of each coordinate and discards her multiplicities and the original system; Bob keeps and aligns his payload where $\beta_i = \bar{\alpha}_i$ and otherwise replaces it by a dummy of dimension d_{α_i}; Alice then samples generalized Weyl twirls on every coordinate and announces them; finally hashing (2.2). All messages go from Alice to Bob, and there is no acceptance test on the whole string. Pulling the Bell-failure effect of coordinate i back through these steps gives an effect at most $P^A_{i\lambda}(I - \Pi_i)$ on the type-λ branch, so the conditional failure $t_{i\lambda}$ satisfies $a_{i\lambda}t_{i\lambda} \le q_{i\lambda}$. Conditional on α the twirled payload is diagonal in the tensor product of the generalized Bell bases, possibly with correlated error labels, and Bob's marginal is maximally mixed, of dimension $\prod_i d_{\alpha_i}$; including the identical classical registers α on both sides does not change the coherent information, so it is at least $\sum_i \log d_{\alpha_i} - \sum_i H(\text{Bell error}_i \mid \alpha)$. Jensen's inequality with the concave envelope F_d and $a_{i\lambda}t_{i\lambda} \le q_{i\lambda}$ proves the displayed bound; discarding a branch is a decision of Alice that depends only on her local label. For the last claim, transferring a length-two product to the opposite party costs 2μ, so Alice's cross-commutators are at most $\eta + 4\mu$, and $\sqrt{2}\mu \le \kappa$ by Lemma 2.11(d). □

One-dimensional sectors contribute nothing even when they carry nontrivial characters; the resource is the weight on the noncommutative blocks. Exact irrep flags and their maximally entangled factors are established [102, 68]; the contribution here is the per-coordinate mismatch accounting on an arbitrary original state.

Corollary 6.8 (Three-cycle visibility)**.** *For $G = S_3$, suppose the hypotheses of Theorem 6.7 hold with Bob cross-commutator bound η and mirror bound μ, put $\kappa = \eta + 4\mu$, and suppose a three-cycle c has $\|(u_i^A(c) - I)\Psi\| \ge \Delta$ for every i. If $N\kappa^2 \le \Delta^2/38400$ then $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge \Delta^2 N/16$.*

Proof. The irreps of S_3 have dimensions $1, 1, 2$, and $2I - R(c) - R(c)^* = 3P_{\mathrm{std}}$ in the regular representation. Hence $a_i := a_{i,\mathrm{std}} = \frac{1}{3}\|(R_i^A(c) - I)\omega\|^2 \ge \frac{1}{3}(\Delta - \sqrt{20N\kappa})_+^2 \ge \Delta^2/12$, and $q_i/a_i \le 50N\kappa^2 \cdot 12/\Delta^2 \le \frac{1}{64}$. For $t \le \frac{1}{2}$ the function $a \mapsto a[1 - F_2(q/a)]$ increases in a (its derivative is $1 + \log(1-t) \ge 0$ at $t = q/a$) and decreases in q, so each summand of Theorem 6.7 is at least $\frac{\Delta^2}{12}[1 - h(\frac{1}{64}) - \frac{\log 3}{64}] \ge \frac{3}{4} \cdot \frac{\Delta^2}{12}$. □

The exactness of each individual copy is supplied, from approximate data, by a local rounding.

Lemma 6.9 (Local S_3 rounding)**.** *Let S_i^P, T_i^P be reflections on each party with $\|(S_i^A - S_i^B)\Psi\|, \|(T_i^A - T_i^B)\Psi\| \le \mu_0$, $\|((S_i^P T_i^P)^3 - I)\Psi\| \le e$, and $\|[Q_i^B, Q_j^B]\Psi\| \le \eta_0$ for $i \ne j$ and $Q \in \{S, T\}$. There are exact local S_3 representations u_i^P with $u_i^P(s) = S_i^P$ satisfying the hypotheses of Corollary 6.8 with $\kappa = 9\eta_0 + 144\mu_0 + 36e$; if Bob's raw three-cycle has $\|(S_i^B T_i^B - I)\Psi\| \ge \Delta_0$, Alice's rounded three-cycle has displacement at least $\Delta = \Delta_0 - 2\mu_0 - e$.*

Proof. Let $C = S_i T_i$ on either party, so $SCS = C^*$. Let r_3 send each point of the unit circle other than -1 to a nearest cube root of unity, choosing $r_3(e^{\pm i\pi/3}) = 1$ at those ties, and set $r_3(-1) = 1$. Then $r_3(\bar{z}) = \overline{r_3(z)}$, and $|r_3(z) - z| \le |z^3 - 1|$ on the whole circle (at $z = -1$ both sides equal 2). Put $C_0 = r_3(C)$, $u(s) = S$ and $u(t) = SC_0$. Then $C_0^3 = I$ and $SC_0S = r_3(C^*) = C_0^*$, so $u(s), u(t)$ are reflections with $(u(s)u(t))^3 = I$: an exact S_3 representation. Functional calculus gives $\|(C_0 - C)\Psi\| \le e$ and $\|(u(t) - T)\Psi\| \le e$. Every element of S_3 is a word of length at most three in s, t, and for involutive generators the reversed word represents the inverse, so the mirrors (6.10) hold with $\mu = 3\mu_0 + 6e$. Replacing the letters of a word of length $L \le 6$ by rounded ones, with raw suffixes moved to the other party, costs at most $Le + L(L-1)\mu_0 \le 6e + 30\mu_0$; two normal forms from different copies

commute after at most nine raw swaps with suffixes of length at most four, each costing $\eta_0 + 8\mu_0$. So Bob's cross-commutators are at most $9\eta_0 + 132\mu_0 + 12e$, and $\kappa = \eta + 4\mu$ gives the stated value. Finally Bob's raw $S^B T^B$ transfers to Alice's $T^A S^A = (S^A T^A)^*$, which has the same displacement as $S^A T^A$, at cost $2\mu_0$, and rounding costs e. □

6.5 Alice-local recovery and the forward direction

Theorem 6.10 (Alice-local recovery)**.** *Let ρ_{AB} be a normal state with purification $\Psi = \sum_k \sqrt{\lambda_k}|a_k\rangle_A|u_k\rangle_{BE}$. Let Alice apply the* linear *isometry $V_A|a_k\rangle = |\bar{u}_k\rangle_{\bar{B}\bar{E}}$ into fresh registers, where the bar denotes complex conjugation of coordinates in fixed product bases, followed if desired by discarding $\bar{E}$.*

(a) $(V_A \otimes I)\Psi = |\sqrt{R}\rangle$ with $R = \rho_{BE}$, in the vectorization $\bar{B}\bar{E} : BE$. Before the discard both directional rates are unchanged; afterwards neither increases. Bob's marginal and operators are untouched.

(b) If U is a unitary on Bob with an actual unitary Alice mirror of error $\delta \le \sqrt{2}$, then the exact transpose U^T on $\bar{B}$ mirrors U on the output with error at most

$$\phi(\delta) = \sqrt{2\delta^2 - \delta^4/2} \;\le\; \sqrt{2}\,\delta.$$

In binary form, a disagreement probability $p \le \frac{1}{2}$ becomes at most $2p(1-p)$.

(c) If Bob has an exact tensor factor of M qubits with Paulis X_i, Z_i, and Alice has actual unitary predictors with disagreement probabilities $p_{X_i}, p_{Z_i} \le \frac{1}{2}$, then

$$E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \;\ge\; M - \sum_i [h(2p_{X_i}(1-p_{X_i})) + h(2p_{Z_i}(1-p_{Z_i}))],$$

and $E_{\mathrm{D}}^{B\to A}(\rho_{AB}) \ge M - \sum_i [h(p_{X_i}) + h(p_{Z_i})]$.

Proof. (a) The conjugate vectors $\bar{u}_k$ are orthonormal, so V_A is an isometry on the support, extended by a local orthogonal register; it is linear, not an antiunitary operation on the input. On Ψ it gives $\sum_k \sqrt{\lambda_k}|\bar{u}_k\rangle|u_k\rangle = |\sqrt{R}\rangle$. Local isometries preserve, and local channels do not increase, the directional rates; nothing acts on B or E.

(b) Let V be Alice's mirror. The vectors $V\Psi$ and $U\Psi$ purify R and $S = URU^*$ on BE, and $\Re\langle V\Psi, U\Psi\rangle = 1 - \delta^2/2$, so by Uhlmann's transition-probability theorem [139] the root fidelity satisfies $F(R,S) = \|\sqrt{R}\sqrt{S}\|_1 \ge 1 - \delta^2/2$. The affinity $A(R,S) = \operatorname{Tr}\sqrt{R}\sqrt{S}$ satisfies $F^2 \le A$: for a unitary W, $\operatorname{Tr}\sqrt{R}\sqrt{S}W = \operatorname{Tr}(R^{1/4}S^{1/4})(S^{1/4}WR^{1/4})$, whose first factor has squared Hilbert–Schmidt norm A and second at most $\operatorname{Tr}\sqrt{R}W^*\sqrt{S}W \le 1$; maximize over W [73, (44)]. On $|\sqrt{R}\rangle$, $\|(U^T\otimes I - I\otimes U)|\sqrt{R}\rangle\|^2 = \|[U,\sqrt{R}]\|_2^2 = 2 - 2A(R,S) \le 2 - 2(1-\delta^2/2)^2 = 2\delta^2 - \delta^4/2$. For reflections the disagreement probability is a quarter of the squared mirror error, which gives the binary form. Normal infinite systems follow by trace-class approximation.

(c) Bob first discards the rest of his system into the environment; then $\bar{B}$ is exactly M qubits on which Alice holds exact X_i^T, Z_i^T, predicted by Bob's X_i, Z_i with disagreements $p' \le 2p(1-p)$ by (b). Fano's inequality for each bit and subadditivity give $H(X_{\bar{B}}^M \mid B) + H(Z_{\bar{B}}^M \mid B) \le \sum_i [h(p'_{X_i}) + h(p'_{Z_i})]$, and the uncertainty relation with quantum memory [16] gives $H(X^M \mid B) + H(Z^M \mid B) \ge M + S(\bar{B} \mid B)$. Hence $-S(\bar{B} \mid B)$ is at least the displayed bound, and hashing with sender $\bar{B}$, held by Alice, makes it achievable from Alice to Bob. For the reverse statement no lift is needed: Alice measures the binary POVM with observable $\frac{1}{2}(M + M^*)$ of her predictor M, whose disagreement with Bob's reflection Q is exactly $\frac{1}{4}\|(M - Q)\Psi\|^2$; the same Fano, subadditivity and uncertainty argument on Bob's exact register gives $-S(B_0 \mid A)$, achievable from Bob to Alice. If Alice's system is infinite, apply the argument after finite-rank erasure truncations on Alice, which change the finitely many prediction probabilities arbitrarily little. □

The map V_A, followed by the discard of $\bar{E}$, is a Petz recovery map [120], the reversal operation of Barnum and Knill [9], in the pretty-good recovery form of [17, App. A.1], with sides and conventions made explicit; its finite Choi matrix is $(\rho_A^{-T/2} \otimes I)\overline{\rho_{AB}}(\rho_A^{-T/2} \otimes I)$ on the support, and the binary bound $p' \le 2p(1-p)$ also follows from [123, (12)]. We claim no new recovery map. One-shot bounds for one-way distillation with a Petz decoder, through binary hypothesis testing, are given by Can and Wilde [26]; they concern a given state, not the near-winners of a fixed game. Its use here is that the *original* state and Bob's operators are kept, which gives the physical forward direction. Applied to all generators, Alice's lifted operators $D_g = \overline{U_g} = (U_g^*)^T$ form a homomorphism, so the whole exact source, with $J = -I$, is transported to Alice; this is used in Sections 8 and 14.

6.6 Exact synchrony gives accessible maximally entangled factors

Theorem 6.11 (Exact synchrony)**.** *Let G be a finite synchronous game, A_{xa}, B_{xa} finite-dimensional projective measurements and ρ_{AB} a mixed state with $\operatorname{Tr}\rho(A_{xa}\otimes I - I\otimes B_{xa})^2 = 0$ for every x, a. Restricted to the marginal supports, decompose the two measurement algebras into matching irreducible blocks of dimensions d_λ with multiplicity spaces. Then every vector in the support of ρ lies in $\bigoplus_\lambda \Phi_{d_\lambda} \otimes (M_\lambda^A \otimes M_\lambda^B)$, measuring the local central projections gives both parties the same label λ, and the conditional state is exactly $\Phi_{d_\lambda} \otimes \tau_\lambda$, even with coherence between different λ in ρ. If G has a pure Schmidt-rank lower bound $c\ell^{-\alpha}$ at small loss ℓ, then for every fixed $0 < \eta < 1$ an exactly synchronous mixed strategy of loss ε yields, with success at least $1-\eta$, an exact maximally entangled state of rank at least $c_\eta \varepsilon^{-\alpha}$.*

Proof. Exact synchrony means $(A_{xa} \otimes I)v = (I \otimes B_{xa})v$ for every v in the support of ρ; the supports of the marginals are invariant under the local measurements. Vectorizing v gives an operator D with $A_{xa}D = DB_{xa}^T$, an intertwiner of the Alice representation with the transpose of Bob's. By Schur's lemma the space of such intertwiners is $\bigoplus_\lambda |\Phi_{d_\lambda}\rangle \otimes (M_\lambda^A \otimes M_\lambda^B)$. The central projections lie in the local measurement algebras, and measuring them gives the same λ on both sides with probability p_λ, leaving $\Phi_{d_\lambda} \otimes \tau_\lambda$. All game measurements act on V_λ only, so the strategy with Φ_{d_λ} has loss ℓ_λ with $\sum_\lambda p_\lambda \ell_\lambda \le \varepsilon$. The sectors with $\ell_\lambda \le \varepsilon/\eta$ carry weight at least $1-\eta$ and have $d_\lambda \ge c\eta^\alpha \varepsilon^{-\alpha}$; on them Nielsen's majorization theorem [109] converts Φ_{d_λ} deterministically to one common Φ_M. Measuring λ, rejecting the other sectors and discarding multiplicities is a local protocol. □

This is standard finite-dimensional structure [118]; we record it because the label is physically accessible, unlike the label of a convex-roof decomposition. Approximate synchrony does not suffice for an exact output (Remark 15.2), and native constraint-variable games are not intrinsically synchronous: mirroring one representative bit per variable does not give the hypothesis on all row measurements.

6.7 The optimal pairwise budget at small error

Theorem 6.12 (The small-error pairwise budget)**.** *Define $E(M,\kappa)$ to be the infimum of $E_{\mathrm{D}}^{A\to B}(\rho_{AB})$ over finite-dimensional original mixed states and two-party arrays of M nominal Pauli pairs of reflections, with all signed pair defects on each party and all actual mirror errors, measured on a purification of ρ, at most κ. The signs are anticommuting within a pair and commuting across distinct pairs. One can also include the square-root-state commutator bounds explicitly; the conclusion is the same up to constants. All shared flag registers belong to the original state. Then for integers $M \ge 1$ and $0 < \kappa \le 1/100$,*

$$\frac{1}{40000}\min\{M, \kappa^{-2}\} \le E(M,\kappa) \le 5\min\{M, \kappa^{-2}\}.$$

At $\kappa = 0$ the exact value is M. The constants are not claimed sharp. This is the precise optimal-scaling interpretation of the $M\kappa^2$ extraction budget.

Proof. Actual mirrors give square-root-state commutators at most $\sqrt{2}\kappa$ by mirror transfer. Retain any

$$N = \min\{M, \lfloor (10000\kappa^2)^{-1} \rfloor\}$$

pairs. Apply Theorem 6.4 with uniform bound $\sqrt{2}\kappa$; its Bell failure bound is at most $50N\kappa^2 \leq 1/200$. Thus $E_{\mathrm{D}}^{A\to B} \geq N[1 - 2h_2(1/200)] \geq N/2$. Since $\kappa \leq 1/100$, the floored quantity is at least one and $N \geq \min\{M, \kappa^{-2}\}/20000$. This proves the lower constant on the original mixed state.

For the upper bound, M actual Bell pairs give $E \leq M$. For any integer $K \geq 1$, give both players a classical flag $f : [M] \to [K]$, uniformly distributed over all K^M functions, and K Bell pairs:

$$\rho_{M,K} = K^{-M} \sum_f |f, f\rangle\langle f, f| \otimes \Phi_2^{\otimes K}.$$

Use controlled $X_j = X_{f(j)}$, $Z_j = Z_{f(j)}$, with transposes on Alice. Same-pair anticommutation and mirrors are exact. For $j \neq l$, X_j and Z_l commute unless $f(j) = f(l)$, an event of probability $1/K$. A collision has additive-commutator squared norm four, so the original-state weighted squared norm is $4/K$. All other cross types commute. Marginals are scalar within every flag block, so square-root-state commutators vanish.

The original flagged state's coherent information is K: the classical entropy $M \log K$ cancels between Bob and the joint state. Its pure decomposition has average entanglement K, so formation and hashing bounds agree at $E_{\mathrm{D}}^{A\to B} = K$. This counts all physical shared registers without treating their classical flags as entanglement. Choose $K = \lceil 4/\kappa^2 \rceil$, giving $E \leq \min\{M, \lceil 4/\kappa^2 \rceil\} \leq 5 \min\{M, \kappa^{-2}\}$. Exact extraction proves $E(M, 0) = M$, attained by Bell pairs. □

The small-error restriction is essential. At $M = 1$, $\kappa = 2$, take one-dimensional local spaces, a product state and all nominal X, Z equal to I. Mirrors and state commutators vanish; the same-pair anticommutator defect is exactly two. Its rate is zero although $\min\{1, \kappa^{-2}\} = 1/4$. The example extends to larger M. The lower estimate uses the existing extraction theorem. The flagged construction is related to the Pauli-hash models of [101, Proposition 9.4], to the operator-norm overlapping-qubit constructions of Chao, Reichardt, Sutherland and Vidick [28], and to their subsequent state-dependent self-testing work [29]. Their operator-norm construction places n qubits with pairwise overlap at most ε in dimension $n^{O(1/\varepsilon^2)}$; here the defects are state-weighted, and the cost $\lceil 4/\kappa^2 \rceil$ does not depend on M. The resource is the directional distillable entanglement of the original mixed state.

II. The law

7 The law: unreachable perfection forces usable entanglement

This section proves Theorem A of the introduction and three further results: effective certificates for every fixed target, undecidability of the dichotomy across games, and a finite protocol in a coherent white-box model. The certificates and the protocol are proved for rational games in class (S) and rational binary games in class (C).

7.1 The two classes

Definition 7.1 (The classes of Theorem A)**.** A finite game G, with a fixed question distribution μ, belongs to class (S) or class (C) as follows.

(S) *Synchronous, with tested diagonals.* Both players have the same finite question and answer sets, unequal answers lose on equal questions, and every diagonal pair (x, x) has $\mu(x, x) > 0$. The relations imposed are the vanishing of $e_{xa}e_{yb}$ for the losing answer pairs (a, b) on question pairs with $\mu(x, y) > 0$ (the *supported losing-pair quotient*). A predicate merely called synchronous whose diagonals have zero weight is not in the class.

(C) *Constraint-variable consistency.* There are finitely many variables with finite domains and finitely many constraints. Alice receives a constraint and answers a satisfying assignment to its variables; Bob receives one variable of that constraint and answers a value; they win if the values agree. Every constraint-variable incidence of every stated constraint has positive probability. A variable that occurs in no constraint is never asked; we discard such variables, which changes neither the correlations nor the demand, so every variable lies in some constraint. Native finite binary constraint-system games, including linear system games, are the main examples.

In both classes we assume that the finite-dimensional quantum value is one, so the set in (2.5) is nonempty for every $\varepsilon > 0$.

For class (C) let $\mathcal{A}(G)$ be the universal unital C^*-algebra generated by a projection-valued measure for each variable, with the variable projections within each constraint commuting and every ordered product of projections representing a forbidden assignment equal to zero. A finite-dimensional perfect strategy exists if and only if $\mathcal{A}(G)$ has a nonzero finite-dimensional representation: given one, Bob measures the variable PVMs, Alice measures the transpose of the joint PVM of the queried constraint, and the state is Φ_d; the converse is Remark 16.2 below in its finite-dimensional case. For class (S) the same holds with the supported losing-pair quotient, with Alice using transposes of the question PVMs.

7.2 Theorem A

Theorem 7.2 (Theorem A)**.** *Let G be a finite game in class (S) or class (C). Then*

$$G \text{ has no finite-dimensional perfect strategy} \quad \Longleftrightarrow \quad d_G^{\rightarrow}(\varepsilon) \longrightarrow \infty \quad (\varepsilon \downarrow 0).$$

More precisely, if G has no finite-dimensional perfect strategy, then for every integer $r \geq 1$ and every $\gamma > 0$ there is $\varepsilon_0(G, r, \gamma) > 0$ such that every finite-dimensional mixed strategy of loss at most ε_0 admits product local isometries extracting r Bell pairs with fidelity at least $1-\gamma$; on any purification the output is within $\sqrt{2\gamma}$, in vector norm, of $\Phi_{2^r} \otimes \chi$, all remaining systems, the environment included, being in the junk χ. The isometries may depend on the strategy and need not be efficient. The threshold is for fixed r and γ; no rate in r is asserted.

The direction "$\Leftarrow$" is immediate: a finite-dimensional perfect strategy is admissible at every ε, and its directional rates are at most the logarithm of its smaller local dimension. The same remark shows that the extraction clause needs its hypothesis: a game with one question and one legal answer is won by a product state. The substance is "$\Rightarrow$", which we prove in four steps: a tracial limit of Bob's moments, the exclusion of every finite type I summand, bounded polynomial approximation of Pauli matrices in the limit, and evaluation of those polynomials back on the original strategies.

Proof of Theorem 7.2, class (S). Step 1: actual mirrors and a tracial limit. Let (ρ_n) be finite-dimensional strategies of losses $\varepsilon_n \to 0$. Dilate locally to projective measurements, and let ψ_n purify the dilated state. For a question x, $\sum_a \|(P^A_{xa} - P^B_{xa})\psi_n\|^2 = 2\Pr[a \neq b \mid (x,x)] \leq 2\varepsilon_n/\mu(x,x)$, so every generator has an actual mirror with a common error $\delta_n \to 0$. Let $\mathcal{A}_0$ be the universal unital C^*-algebra of the free PVMs on the finitely many questions, separable, and $\tau_n(F) = \langle \psi_n, F(P^B)\psi_n\rangle$. After passing to a subsequence, $\tau_n \to \tau$ weak-$*$. Write θ for word reversal followed by substitution of Alice's PVMs; it extends to an isometric $*$-antiautomorphism of $\mathcal{A}_0$ followed by a representation, and (6.1) holds. For a generator e and a word w, $\tau_n(ew) = \langle e^B\psi_n, w^B\psi_n\rangle$, and replacing $e^B\psi_n$ by $e^A\psi_n$, commuting e^A past w^B and replacing it back costs at most $2\delta_n$. So $|\tau_n(ew) - \tau_n(we)| \leq 2\delta_n$, and by iteration, linearity and density τ is a trace. If (a,b) loses at (x,y) with $\mu(x,y) > 0$, then $\tau_n(e_{xa}e_{yb})$ is, up to δ_n, the actual probability of (a,b) at (x,y), which tends to zero; so $\tau(e_{xa}e_{yb}) = 0$ and, τ being a trace, $\|e_{xa}e_{yb}\|^2_{2,\tau} = \tau(e_{xa}e_{yb}) = 0$. The tracial null space is a two-sided ideal, so in the GNS representation π_τ every forbidden product vanishes as an operator. Put $\mathcal{M} = \pi_\tau(\mathcal{A}_0)''$, a finite von Neumann algebra with faithful normal trace t on a separable space.

Step 2: no finite type I summand. Suppose $\mathcal{M}$ had a nonzero type I central summand. By the structure theorem it contains a nonzero homogeneous piece $L^\infty(X,\nu) \otimes M_d(\mathbb{C})$ for some finite d; there is no type I_∞ piece in a finite algebra, and separability gives standard measure spaces. Represent the finitely many generating projections on this piece by measurable fields of $d \times d$ matrices. The finitely many defining relations (projections, PVM sums, forbidden products) hold almost everywhere, hence on one common conull set; at any point of it the matrices are PVMs satisfying all relations, and with Alice's transposes and Φ_d they give a finite-dimensional perfect strategy. This also excludes diffuse type I centres. So, if G has no finite-dimensional perfect strategy, $\mathcal{M}$ is of type II, not necessarily a factor, and for every d it contains a unital copy of $M_d(\mathbb{C})$: the centre-valued trace divides 1 into d equivalent projections [27, Lemma 1.4].

Step 3: bounded approximation. Fix r, put $m = 2r$, and let $q_1, \dots, q_m \in \mathcal{M}$ be the standard Pauli reflections $X_1, \dots, X_r, Z_1, \dots, Z_r$ of a unital $M_{2^r}(\mathbb{C})$, with $q_jq_k = s_{jk}q_kq_j$. Fix $\eta \in (0,1)$. By Kaplansky's density theorem [79] there are self-adjoint contractions $a_j \in \pi_\tau(\mathcal{A}_0)$ with $\|a_j - q_j\|_{2,t} < \eta/2$. Each lifts to a self-adjoint contraction $h_j \in \mathcal{A}_0$ (lift, average with the adjoint, and apply the clipping function $\max(-1, \min(t,1))$, which commutes with the quotient map); approximate h_j in norm by a self-adjoint $*$-polynomial F_j within $\eta/2$. Then $\|F_j\|_{\mathcal{A}_0} \leq 2$, $\|\pi_\tau(F_j) - q_j\|_{2,t} < \eta$, and, the trace making bounded multiplication continuous in $\|\cdot\|_{2,t}$,

$$\|\pi_\tau(F_j^2 - 1)\|_{2,t} \leq 3\eta, \qquad \|\pi_\tau(F_jF_k - s_{jk}F_kF_j)\|_{2,t} \leq 6\eta. \tag{7.1}$$

Step 4: back on the original strategies. Fix these polynomials before n grows, and put $H^B_j = F_j(P^B_n)$ and $H^A_j = \theta(F_j)(P^A_n)$, self-adjoint of norm at most two on both parties. Moment convergence turns the squared vector norms of the fixed defect polynomials on ψ_n into their τ-norms, so Bob's square and pair defects are eventually at most $3\eta + o(1)$ and $6\eta + o(1)$. By (6.1) the mirrors $\|(H^A_j - H^B_j)\psi_n\|$ tend to zero, and since $\theta(F_jF_k - s_{jk}F_kF_j) = -s_{jk}(\theta F_j \theta F_k - s_{jk}\theta F_k \theta F_j)$ the same bounds hold for Alice. Round locally, $Q^P_j = \operatorname{sgn}(H^P_j)$ with $\operatorname{sgn}(0) = 1$; the scalar inequality $|\operatorname{sgn} t - t| \leq |t^2 - 1|$ bounds the rounding error on ψ_n by the square defect. Inside a product, $\|(Q_j - H_j)H_k\psi_n\| \leq \|H^{\mathrm{op}}_k(Q_j - H_j)\psi_n\| + \|Q_j - H_j\|\|(H_k - H^{\mathrm{op}}_k)\psi_n\|$, where H^{op}_k is the actual opposite-party mirror; so each product replacement costs at most three times the rounding error plus three times the mirror error. Altogether the rounded reflections satisfy, on both parties,

$$\|(Q^P_jQ^P_k - s_{jk}Q^P_kQ^P_j)\psi_n\| \leq \kappa_n, \quad \|(Q^A_j - Q^B_j)\psi_n\| \leq \kappa_n, \quad \limsup_n \kappa_n \leq 100\eta.$$

Lemma 6.5 now gives product local isometries whose output has singlet infidelity at most $2r(2C_r+1)^2\kappa_n^2/4$ and, by (2.3), $E^{A\to B}_{\mathrm{D}}(\rho_n) \geq r[1 - 2h(p)]$ with $p = (2C_r+1)^2\kappa_n^2/4$. Letting

$\eta \to 0$ after $n \to \infty$, $\liminf_n E_{\mathrm{D}}^{A\to B}(\rho_n) \geq r$, and for every γ the extraction has fidelity at least $1-\gamma$ for all large n.

Uniformity. If for some finite K there were strategies with losses tending to zero and $E_{\mathrm{D}}^{A\to B} \leq K$, apply the above to that sequence with $r > K$: a contradiction. If for some r, γ no threshold worked, there would be a sequence with losses tending to zero none of which admits an r-pair extraction of fidelity $1-\gamma$: again a contradiction. The vector-norm form follows from the fidelity form by projecting onto the Bell factor. □

Proof of Theorem 7.2, class (C). After local projective dilation let $B_{j,a}$ be Bob's variable-value projections and $C_{i,j,a}$ Alice's coarse-grained projections of the joint assignment for constraint $i \ni j$. For fixed i the $C_{i,j,a}$ commute and their forbidden products vanish exactly, since Alice's answers are satisfying assignments. Positive probability of every incidence and vanishing loss give $\|(C_{i,j,a} - B_{j,a})\psi_n\| \to 0$. Choose once an incident constraint $i(j)$ for each variable and use $A_{j,a} = C_{i(j),j,a}$ as the actual Alice mirrors of the universal Bob generators; distinct proxies need not commute. Step 1 now runs on the universal free-PVM algebra of the variable measurements and gives a trace. A Bob word on the variables of one constraint i transfers to the reversed word in that constraint's own commuting $C_{i,j,a}$, so Bob's within-constraint commutators and forbidden products have vector residuals tending to zero; their squared τ-norms vanish and the tracial GNS ideal makes them exact operator relations in $\mathcal{M}$. A finite-dimensional representation of this constraint algebra would be a finite-dimensional perfect strategy, so Steps 2–4 apply verbatim. □

Remark 7.3 (What is used, and what is not). The limit is a tracial GNS representation of finitely many approximate algebra elements, not an assumed spatial perfect strategy and not a matrix ultraproduct with an unproved lift; every operation is evaluated back on the original strategies, with their purifying environments. The exact tracial correspondence for perfect consistency is that of Cleve–Mittal and of Paulsen and collaborators [34, 118, 83]; approximate tracial rounding of near-perfect strategies is due to Vidick, Marrakchi–de la Salle, Paddock and Culf [145, 97, 112, 45]. Rounding to a favorable tracial strategy is not by itself a lower bound on E_{D} of an arbitrary original mixed state, and that is the conclusion here. Approximating a strategy blockwise by maximally entangled ones would give a different state on each block, while extraction needs one pair of local isometries on the original state, a difficulty discussed by Vernooij and Zhao [143, Section 1.2]; the proof above avoids it, because the polynomials are evaluated on the original strategies themselves. Nor does the proof use amenability of the limit trace: amenability need not pass from a trace on a universal algebra to a quotient [149, Remark 5.15], but the limit here is only a tracial state of the free algebra $\mathcal{A}_0$, the losing products vanish in its GNS representation, and the polynomials are chosen in $\mathcal{A}_0$, where the moments converge. The type dichotomy is the one behind the infinite-dimensional results compared in Section 1.2: Keyl, Matsui, Schlingemann and Werner prove the infinite-one-copy-entanglement statement for every normal joint state of hyperfinite factors in Haag duality that are not of type I [81, Theorem 4.9], and van Luijk, Stottmeister, Werner and Wilming remove hyperfiniteness for extraction from pure vector states [142, Theorem 54]; the mixed-state assertion for finite strategies is proved here directly. No classification of arbitrary games is proved, and whether the conclusion extends beyond classes (S) and (C) is open (Section 1.6).

Remark 7.4 (Finite menus). For fixed (G, r, γ), finitely many polynomial tuples suffice for all sufficiently good strategies. The tracial states of the game algebra with zero losing products form a compact set T_G; each $\tau \in T_G$ has a tuple whose defects (7.1) are below a fixed tolerance, which is a weak-$*$ open condition; compactness gives a finite cover. If a sequence of strategies with losses tending to zero defeated every tuple of the cover, a moment subsequence would

converge into T_G and into one of the open sets, a contradiction. The successful tuple may depend on the strategy.

7.3 Effective certificates

The compactness proof gives no threshold. For rational games, with binary variables in class (C), a threshold can nevertheless be computed, by a search that is guaranteed to halt under the promise.

Theorem 7.5 (Effective fixed-target certificates)**.** *Let G be a rational game in class (S), or a rational finite binary constraint-variable game in class (C), promised to have finite-dimensional quantum value one and no finite-dimensional perfect strategy. For fixed r and rational $0 < \gamma < 1$, a partial algorithm halts and returns a rational $\varepsilon_0 > 0$, a finite menu of rational universally contractive self-adjoint polynomial arrays, each followed by local sign rounding and the fixed signed regular isometries of Lemma 6.5, and a finite rational tracial-moment infeasibility certificate, such that every strategy of loss at most ε_0 is decoded by some member of the menu with an r-pair output of fidelity at least $1-\gamma$. No useful running time, closed-form dependence on r, or decision of the promise is asserted.*

Proof. Universally contractive rational polynomials. Enumerate rational word polynomials q and put $p = \frac{1}{2}(q+q^*)$, $C = 1 + \sum_w (|\Re c_w| + |\Im c_w|)$, so $\|p\| \le C$ in every PVM representation. With $c(t) = \max(-1, \min(t, 1))$ and the Bernstein polynomial $b_{C,n}(t) = \sum_{k=0}^{n} c(-C + 2Ck/n)\binom{n}{k} u^k (1-u)^{n-k}$, $u = (t+C)/(2C)$, the polynomial $F = b_{C,n}(p)$ has rational coefficients and satisfies $F = F^*$ and $\|F\| \le 1$ exactly in every representation, because the Bernstein weights are commuting positive operators summing to one. Also $\sup_{|t|\le C} |b_{C,n}(t) - c(t)| \le C/\sqrt{n}$, and these F are norm-dense in the self-adjoint unit ball of $\mathcal{A}_0$.

A finite cover. Put $m = 2r$, $C_r = m(1+2m)$ and the rational target $a = \gamma/[10^8 (r+1)(2C_r+1)^2]$. For an array $F = (F_1, \dots, F_m)$ let $d_F = \sum_j (F_j^2 - 1)^2 + \sum_{j<k} (F_j F_k - s_{jk} F_k F_j)^* (F_j F_k - s_{jk} F_k F_j)$. By Steps 2–3 of the proof of Theorem 7.2, every perfect trace $\tau \in T_G$ has an enumerated array with $\tau(d_F) < a/2$, and compactness gives a finite menu $F^{(1)}, \dots, F^{(K)}$ with $T_G \cap \{\tau(d_{F^{(k)}}) \ge a \ \forall k\} = \emptyset$.

A complete moment hierarchy, with two-sided PVM relations. At level N use moments y_w for words of length at most $2N$, with $y_1 = 1$, $y_{w^*} = \overline{y_w}$, a positive semidefinite moment matrix $[y_{u^* v}]_{|u|,|v| \le N}$, the relations $y(u(e^2 - e)v) = 0$, $y(u(e - e^*)v) = 0$ and $y(u(\sum_a e_{xa} - 1)v) = 0$ for *all* left and right words u, v within the degree budget, cyclicity $y_{uv} = y_{vu}$ within the budget, $y(e_{xa} e_{yb}) = 0$ for supported losing pairs (for class (C): $y(Q^*Q) = 0$ for the within-constraint commutators and forbidden products Q), boxes $|\Re y_w|, |\Im y_w| \le 1$, and $y(d_{F^{(k)}}) \ge a$ for every k. This is a finite closed rational semialgebraic system. If it were feasible at every level, a diagonal limit would be a positive normalized tracial functional; the boundedness of the generators in its GNS construction uses $y(v^*(e^2 - e)v) = 0$, an instance of the two-sided relations, which gives $\|e[v]\|^2 = \langle [v], e[v] \rangle$ and hence $\|e[v]\| \le \|[v]\|$. The limit would lie in T_G and be bad for the whole menu, contradicting the cover. So some finite level is infeasible, and by compactness it stays infeasible after relaxing only the trace and game equalities to absolute error ξ, for some rational $\xi > 0$. Exact real quantifier elimination decides each such instance [35].

Transfer to near-winners. Let $\mu_{\min}$ be the smallest positive incidence or diagonal weight; a strategy of loss ε has generator mirrors $\delta \le \sqrt{2\varepsilon/\mu_{\min}}$. Its Bob moments satisfy the PVM relations, positivity and boxes exactly, the cyclic equalities up to $4N\delta$, and the game equalities up to $\varepsilon/\mu_{\min} + \delta$ (for class (C), $\tau_B(Q^*Q) \le B_0^2 \delta^2$ with $B_0 = \max\{4, \max_i |i|\}$, by row transfer). Let $K_0 \ge 1$ be a rational bound on $\sum_w |c_w||w|$ for every polynomial of the menu and its

defects, and put

$$z = \min\{1, a/K_0, \xi/(4N+1), \xi/2\}, \qquad \varepsilon_0 = \min\{1, (\mu_{\min}/2)z^2, \mu_{\min}\xi/2\}$$

(for class (C), $z = \min\{1, a/K_0, \xi/(4N+1), \xi/(B_0+1)\}$ and $\varepsilon_0 = (\mu_{\min}/2)z^2$). For loss at most ε_0 the moments satisfy the relaxed system, so some array has $\tau_B(d_F) < a$: all Bob square and pair residuals are at most $\sqrt{a}$, every raw polynomial mirror is at most $K_0\delta \le a$, and Alice's residuals are at most $2\sqrt{a}$. Sign rounding and product replacement as in Step 4 give a common bound $\kappa = 100\sqrt{a}$, and Lemma 6.5 gives singlet infidelity at most $2r(2C_r+1)^2(100^2 a)/4 < \gamma$.

The search. Enumerate arrays, menus, levels and rational $\xi \to 0$, and run quantifier elimination on each finite relaxed system; under the promise an infeasible candidate appears, and the algorithm outputs it with ε_0. Every returned certificate is sound regardless of the promise. □

Halting for one target does not decide the absence of finite-dimensional perfect strategies: a game with a large finite perfect strategy can also satisfy a small-r certificate. The rational Bernstein construction, tracial moment hierarchy, quantifier elimination, type decomposition and Kaplansky density are established tools; the composed fixed-target certificate for original mixed near-winners is what is new. The moment hierarchy is the tracial form of the semidefinite hierarchies of Navascués, Pironio and Acín for quantum correlations and for noncommutative polynomial optimization [108, 121]; tracial moment matrices are studied by Burgdorf and Klep [24]. The relation $y(u(e^2 - e)v) = 0$ must be imposed for all left and right words; one-sided imposition is not enough for the boundedness step.

7.4 Undecidability of the dichotomy

Theorem 7.6 (Bounded versus divergent demand is undecidable)**.** *There is a fixed bounded-output format, rational binary constraint-variable games with every incidence positively tested, contexts of arity at most four, Alice at most sixteen answer labels and Bob two, such that, among games of this format promised to have finite-dimensional quantum value one:*

(a) the games with bounded demand, $\lim_{\varepsilon\to 0} d_G^{\rightarrow}(\varepsilon) < \infty$, are exactly those with a finite-dimensional perfect strategy, and form a semidecidable set;

(b) the games with divergent demand form an undecidable set that is not semidecidable.

The numbers of questions may grow. The statement concerns finite game descriptions, not the distillability of a given state or the capacity of a channel.

Proof. Inputs. By Slofstra [132, Corollary 1.4], it is undecidable whether a finite binary linear system game has a perfect strategy in finite dimension. By [132, Theorem 1.1] and its construction [132, §4, §5, Proposition 5.1 and Remark 5.5] there is a fixed linear system game H with a perfect strategy in C_{qa} but none in C_{qs}, hence with finite-dimensional quantum value one and no finite-dimensional perfect strategy. That construction already has Alice eight answers and Bob two, with ternary equations, so no further arity reduction of H is needed.

Ternary inputs. An input linear system G is first replaced by one with contexts of at most three variables and the same finite-dimensional operator solutions: for a row $x_1 \cdots x_t = (-1)^b$ add, freshly for this row, z_{ij} with contexts $x_i x_j z_{ij} = 1$ for $i < j$, prefix variables $p_2 = z_{12}$ and $p_{k-1}x_k p_k = 1$, and the unary row $p_t = (-1)^b$. Given a solution put $z_{ij} = x_i x_j$ and $p_k = x_1 \cdots x_k$; conversely the pair contexts force the x_i to commute and the prefixes recover the parity.

The selector. With disjoint variable sets for G and H and one new flag f, let $K(G)$ have a context $C \cup \{f\}$ for each context C of G, allowing exactly the assignments with $f = 1$ or satisfying the constraint of C, and a context $D \cup \{f\}$ for each context D of H, allowing $f = 0$ or the constraint of D; test every incidence, including every occurrence of f, with positive rational probability. The flag occurs in every context, so in any operator solution its projections e_0, e_1 commute with every variable and are central; all original context commutations remain required on both branches.

Value one. Play near-perfect finite strategies of H with the flag deterministically 1 and the G variables set to fixed scalars: G contexts and all flag incidences pass exactly and the loss is a fixed weighted sum of the losses of H. So every $K(G)$ has finite-dimensional value one, whatever G is.

Perfection is preserved exactly. If G has a finite-dimensional operator solution, take $e_0 = I$, keep it, and set the H variables to scalars. Conversely, a finite-dimensional perfect strategy of $K(G)$ gives a finite-dimensional representation of its algebra (Remark 16.2); split it by the central flag; the e_1 part would solve H and is therefore zero, so $e_0 = I$ and the G operators solve G. Thus $K(G)$ has a finite-dimensional perfect strategy if and only if G does, and by Theorem 7.2 this happens if and only if $d^{\rightarrow}_{K(G)}$ stays bounded.

Undecidability and non-semidecidability. A decision procedure for bounded demand on the promised class would decide Slofstra's problem. Finite perfection is semidecidable: for $d = 1, 2, \ldots$ decide by quantifier elimination whether a d-dimensional operator solution exists. If divergence were also semidecidable, running both searches on $K(G)$ would decide Slofstra's problem.

Bounded output. The fallback H is replaced by a fixed ternary H_3 by the same gadget; H_3 still has no finite perfect strategy, and it has value one on the same states: for each auxiliary z or p representing a fixed product w of L original variables of one row, Bob uses $Q_w = \operatorname{sgn}(\frac{1}{2}(U_w + U_w^*))$ with $U_w = \prod_{x \in w} B_x$, and Alice computes the parity bits from her original row assignment; the estimates $\|(U_w - A_w)\psi\|, \|(U_w^* - A_w)\psi\| \le L\delta$, $\|(F_w^2 - I)\psi\| \le 2L\delta$ and $\|(Q_w - A_w)\psi\| \le 3L\delta$ show that every new incidence error tends to zero. With G_3 and H_3 the selector has contexts of at most four variables including the common flag, so Alice's answers are encoded in sixteen labels; invalid labels lose and can be relabelled to valid ones without increasing loss. □

The undecidable algebraic core is Slofstra's; the physical conclusion depends on Theorem 7.2. The selector is an elementary gated disjunction, and it is what keeps the promise: every $K(G)$ has value one through the fixed game H_3, whatever G is, and only the existence of a finite-dimensional perfect strategy depends on G. A closely related disjunction, in which the players choose which of two games to play, is introduced by Mančinska and Schmidt: combining a game of value one without a finite-dimensional perfect strategy, such as Slofstra's, with a synchronous game that self-tests a perfect finite-dimensional strategy, they obtain an exact self-test that is not robust [95, Sections 4 and 5]. Here the choice is a central flag inside a Boolean constraint game, and the fixed nonattaining branch keeps the value-one promise for every input. A gapped problem, value one against value at most $s < 1$, would not serve here, since it does not separate attainment from nonattainment inside the value-one class. For general two-player games such a gapped problem is undecidable by the theorem $\mathrm{MIP}^* = \mathrm{RE}$ of Ji, Natarajan, Vidick, Wright and Yuen, which reduces the halting problem to deciding whether a game has entangled value one or at most one half [75]. Culf and Mastel show that such gapped problems are already RE-complete for constraint-variable games of every Boolean constraint language with NP-complete classical satisfiability, and for 3-colouring [46, Theorem 4.14 and Corollary 4.18]. For example, their 3SAT instances have constraints of arity three and Boolean variable answers, and their 3-colouring instances have binary constraints and a three-colour alphabet. These are value-gap results. Our dichotomy

concerns finite perfect strategies versus diverging usable entanglement under a value-one promise; we do not know whether its arity four and the sixteen Alice labels of Theorem 7.6 can be lowered. The theorem is compatible with Theorem 7.5: on a divergent instance every finite target eventually has a certificate, but no finite prefix of such certificates recognizes divergence. We do not claim that no computable precision modulus exists for a fixed game; a computably specified fixed value-one game has a computable upper envelope of its demand by enumeration of finite witnesses.

7.5 A finite protocol under coherent white-box access

The effective certificate can be compiled into one finite protocol, independent of the hidden state and dimension, if the players have coherent access to their own measurement circuits. This is a coherent white-box oracle model, *not* a device-independent one.

Theorem 7.7 (White-box protocol)**.** *Let G be a rational game in class (S) or a rational binary game in class (C) satisfying the promise of Theorem 7.5, and fix r and rational $0 < \gamma < 1$. Assume: independent identically distributed copies of the original state; fixed memoryless local measurement implementations; physical access to the payload and to reversible Naimark circuits V_x and V_x^* with known outcome labels; pure ancillas; and coherent control of known circuits. A partial compiler returns a loss threshold $\varepsilon_* > 0$ and* one *finite protocol, independent of the hidden state, its dimension and its matrix descriptions, with the following property. On every finite-dimensional mixed strategy of loss at most ε_*, Alice estimates her local defects on sacrificed copies, announces a menu index to Bob, and both apply fixed coherent decoders to one retained copy, which yields r qubits per side with singlet fraction at least $1 - \gamma$; all dirty work registers are kept in the accounting, and the guarantee holds against every purification of the retained input. One sufficient number of sacrificed copies is*

$$K\left\lceil 2\Lambda_D^2 b^{-2} \ln(2K/\zeta)\right\rceil, \qquad b = \frac{\gamma}{10^8(r+1)(2C_r+1)^2},$$
$$\zeta = \frac{\gamma}{8}, \qquad m = 2r, \qquad C_r = m(1+2m),$$

with K the menu size and Λ_D the recorded coefficient bound of the defect polynomials. All communication is from Alice to Bob. The implementation uses a linear combination of unitaries, a two-reflection walk with phase-estimation sign approximation, and the reversible measurement access; the gate costs of the unknown oracles are not bounded.

The proof is in Appendix A. The theorem does not certify near-winning behaviour autonomously: local tests cannot distinguish an entangled input from a product input with the same marginal. Nor does it leave the environment's marginal unchanged after conditioning on the selection transcript; only the retained copy's AB marginal is unchanged, and the bound holds for every conditional purification. A device that offers only classical outputs, or only destructive measurements, does not satisfy the interface: from classical outputs and fresh local ancillas only separable states can be prepared.

Universal distillation from unknown mixed states is known in a fixed finite dimension. Takagi, Watanabe, Matsuura, Arai and Hayashi [135] achieve the distillable entanglement of an arbitrary unknown state through tomography, state merging for compound sources and uniform convergence, with almost one-way communication: Bob sends back tomographic information at a vanishing rate [135, Theorem 8 and Appendix B.3(f)]. Tirone, Mele, Giovannetti and Lami [138] achieve it independently with tomography by local operations and classical communication and a robust adaptive distillation routine. Both build on protocols for compound sources [20] and the earlier work cited there. So the possibility of distilling from an unknown state is not what Theorem 7.7 adds. It instead uses a fixed-game promise and

reversible coherent measurement access to compile a finite dimension-uniform menu, selects a decoder with Alice-local data, and extracts a fixed target from one retained input copy, using only communication from Alice to Bob. It does not optimize a universal asymptotic rate and does not provide an efficient device-independent procedure.

8 Directional invariance

A one-way rate has a direction, and a game could in principle demand usable entanglement in one direction only. This section shows that, for agreement games, it cannot: one Alice-local channel turns every near-winning strategy into a strategy of at most three times the loss whose two directional rates are equal and no larger than either original rate. The channel is a Petz recovery map [120] in pretty-good form, its output is the reflected state of Dutta and Faulkner [58], and the correlation maps below are KMS-adjoint correlation maps, related to the weighted correlation-operator framework of quantum maximal correlation [10]; the theorems of this section are compositions of these known objects, and a bounded search found no preemption of them.

8.1 A local symmetric output that keeps Bob

Theorem 8.1 (Local symmetry with Bob kept)**.** *Let ρ_{AB} be a normal state and let Alice apply the channel of Theorem 6.10, the isometry followed by the discard of $\bar{E}$. Its output $\tau_{\bar{B}B}$ satisfies*

$$\tau_B = \rho_B, \qquad \mathrm{Swap}\,\tau\,\mathrm{Swap} = \bar{\tau}.$$

The two directional rates of τ are equal, and this common rate is at most each *directional rate of ρ_{AB}. Bob's entire operator tuple and his marginal are unchanged, so every specified Bob word or relator error, state commutator, complex phase and visibility is unchanged, including for soft tuples that satisfy no relation exactly. An actual Alice mirror of error $e \le \sqrt{2}$ of a Bob generator becomes the exact transpose with error at most $\phi(e)$, and at most $\sqrt{2}e$ always. Consequently, for any class of sources closed under this local output, the infimum of either directional rate at mirror error e dominates the infimum of the other at $\phi(e)$, all Bob budgets fixed.*

Proof. With $R = \rho_{BE} = \sum_k \lambda_k |u_k\rangle\langle u_k|$, the lifted vector is $|\sqrt{R}\rangle = \sum_k \sqrt{\lambda_k}|\bar{u}_k\rangle|u_k\rangle$, and exchanging the copy $\bar{B}\bar{E}$ with BE sends it to $\sum_k \sqrt{\lambda_k}|u_k\rangle|\bar{u}_k\rangle$, its complex conjugate. Tracing out E and $\bar{E}$ gives $\mathrm{Swap}\,\tau\,\mathrm{Swap} = \bar{\tau}$. Conjugating all Kraus operators of a one-way protocol gives a protocol with the same direction and the same real Bell target, so $E_{\mathrm{D}}^{A\to B}(\bar{\tau}) = E_{\mathrm{D}}^{A\to B}(\tau)$; exchanging the parties reverses the direction. Hence $E_{\mathrm{D}}^{A\to B}(\tau) = E_{\mathrm{D}}^{A\to B}(\mathrm{Swap}\,\bar{\tau}\,\mathrm{Swap}) = E_{\mathrm{D}}^{B\to A}(\tau)$. The channel is Alice-local, so each directional rate of τ is at most the corresponding rate of ρ. The mirror statement is Theorem 6.10(b), and Bob is untouched. □

The output is exactly the reflected state of ρ_{BE} in the sense of Dutta and Faulkner [58]: in the notation of the proof, $\tau = \mathrm{Tr}_{E\bar{E}}\,|\sqrt{R}\rangle\langle\sqrt{R}|$; equivalently, if $\mathcal{N}$ is the channel with $\rho_{AB} = (\mathrm{id}\otimes\mathcal{N})(\psi_{\rho_A})$ for the real symmetric purification $\psi_{\rho_A} = \sum_i \sqrt{\lambda_i}|ii\rangle$, then $\tau = (\bar{\mathcal{N}}\otimes\mathcal{N})(\psi_{\rho_A})$, which makes the symmetry evident. That Alice's recovery map acts, up to complex conjugation, as $\mathcal{N}$ itself is, up to conventions for transposes and the order of the factors, the finite-dimensional case of the correspondence between the Petz recovery map and the swap transpose of quantum couplings [23, Theorem 1]; Parzygnat and Fullwood obtain the same identification as a quantum Bayes' rule [117, (14) and (25)]. What is used here is that this known state is reached by an Alice-local isometry, from the original state, keeping the original Bob and the original environment; the lift is the Petz recovery map [120], in the

pretty-good form of [9, 17, 73, 123]. The output does not preserve a fixed global rank or spectrum, a bound on Alice's dimension smaller than Bob's, or the efficiency of a decoder.

8.2 A universal factor three for agreement games

A bipartite relation $\mathcal{R}$ between Alice's and Bob's answers is *difunctional* if $(a,b),(a',b)$, $(a',b') \in \mathcal{R}$ implies $(a,b') \in \mathcal{R}$; equivalently its bipartite graph is a disjoint union of complete bipartite blocks and isolated vertices, that is, there are label maps l_A, l_B into a common set, with private labels for isolated answers, such that acceptance is exactly $l_A(a) = l_B(b)$. This is standard relational vocabulary [124, 69]. A game is an *agreement game* if its acceptance relation on every positively weighted question pair is difunctional; the label maps may depend on the pair. Projection games, partial projection games, constraint-variable games and constraint-constraint overlap-agreement games (the synchronous BCS games of Kim, Paulsen and Schafhauser [83, Def. 4.2]) are agreement games; synchronous and graph colouring predicates are not, merely by name.

For a finite-dimensional state ρ_{AB} define, on Hermitian operators on the supports of the marginals,

$$S(X) = \rho_B^{-1/2} \operatorname{Tr}_A[(X \otimes I)\rho_{AB}]\rho_B^{-1/2}, \qquad T(Y) = \rho_A^{-1/2} \operatorname{Tr}_B[(I \otimes Y)\rho_{AB}]\rho_A^{-1/2},$$

and the real inner products $\langle X, Y\rangle_\rho = \operatorname{Tr}[\sqrt{\rho}X\sqrt{\rho}Y]$. We use these KMS-adjoint correlation maps, related to the weighted correlation-operator framework of quantum maximal correlation [10], where correlations are measured with the weighted second moments $\operatorname{Tr}(\rho XX^*)$ rather than with this KMS form; we do not identify the two definitions. The positivity, adjointness and contraction needed here are proved below.

Theorem 8.2 (Factor-three normal form)**.** *Let G be a finite agreement game, and take an arbitrary finite-dimensional mixed strategy with legal POVMs $\{P_x^a\}$ for Alice and $\{Q_y^b\}$ for Bob and loss ε. Apply the channel of Theorem 8.1, let Alice use the whole-answer POVM $M_x^a = S(P_x^a)^T$ on $\bar{B}$, and let Bob keep his original tuple. All answers remain legal, every original referee weight is kept, and the new loss satisfies*

$$\varepsilon_{\text{new}} \le 3\varepsilon, \qquad \varepsilon_{\text{new}} \le g(\varepsilon) := 1 - (1-\varepsilon)^3 \quad (0 \le \varepsilon \le \tfrac{1}{2}).$$

There is no dependence on arity, alphabet, game size or smallest question probability. The output has the two equal directional rates of Theorem 8.1, below either original rate. Consequently

$$d_G^{\rightarrow}(\varepsilon) \ge d_G^{\leftarrow}(3\varepsilon), \qquad d_G^{\leftarrow}(\varepsilon) \ge d_G^{\rightarrow}(3\varepsilon),$$

and the same with $g(\varepsilon)$ in its range. This holds also at every fixed number of input copies and fixed Bell-target size, and for one-way secret-key demand when the discarded environments go to the purifying adversary. For an ordinary parallel product of agreement games, applying the channel once to the whole correlated strategy gives the factor three, not 3^m.

Proof. Correlation maps. S and T are positive and unital on the supports, and $\langle S(X), Y\rangle_{\rho_B} = \operatorname{Tr}\rho_{AB}(X \otimes Y) = \langle X, T(Y)\rangle_{\rho_A}$, so they are adjoint in the displayed inner products, which are positive definite on the supports. Hence $K = TS$ is self-adjoint and positive semidefinite; it is also positive and unital, so it contracts the operator norm and its real eigenvalues have modulus at most one. Thus the spectrum of K lies in $[0,1]$ and S, T are contractions. On a direct sum of outcome-indexed copies, $H = \left(\begin{smallmatrix} 0 & T \\ S & 0 \end{smallmatrix}\right)$ is a self-adjoint contraction. For any POVM $\{E_b\}$, $\sum_b \|E_b\|_\rho^2 \le \sum_b \operatorname{Tr}\rho E_b^2 \le \sum_b \operatorname{Tr}\rho E_b = 1$.

The new success. For the channel C of Theorem 8.1, $C^*(R^T) = T(R)$ for Hermitian R on B: with the Choi matrix $J_C = (\rho_A^{-T/2} \otimes I)\overline{\rho_{AB}}(\rho_A^{-T/2} \otimes I)$, $\operatorname{Tr}[C(X)R^T] = \operatorname{Tr}[(X^T \otimes R^T)J_C] =$

$\mathrm{Tr}[\rho_{AB}(\rho_A^{-1/2} X \rho_A^{-1/2} \otimes R)] = \mathrm{Tr}[XT(R)]$. Fix a tested pair (x, y) and coarse-grain both answers into the labels of its difunctional relation: $P_\ell = \sum_{l_A(a)=\ell} P_x^a$ and $Q_\ell = \sum_{l_B(b)=\ell} Q_y^b$, both lists complete, private labels contributing zero on the other side. With $p = (P_\ell)_\ell$, $q = (Q_\ell)_\ell$ and $w = (p, q)/\sqrt{2}$, $\|w\| \le 1$, the original success is $s = \sum_\ell \mathrm{Tr}\,\rho(P_\ell \otimes Q_\ell) = \langle w, Hw\rangle$. Since S is linear, the coarse-grained new effect is exactly $S(P_\ell)^T$, and the new success is

$$s' = \sum_\ell \mathrm{Tr}[\tau(S(P_\ell)^T \otimes Q_\ell)] = \sum_\ell \langle TS(P_\ell), T(Q_\ell)\rangle_{\rho_A} = \langle w, H^3 w\rangle.$$

For $h \in [-1, 1]$, $h^3 - 3h + 2 = (h-1)^2(h+2) \ge 0$, so $s' \ge 3s - 2\|w\|^2 \ge 3s - 2$, that is, the loss at (x, y) is at most three times the original loss. Summing with the original weights gives $\varepsilon_{\mathrm{new}} \le 3\varepsilon$. For the refinement, form the weighted direct sum $w = \bigoplus \sqrt{\mu_{xy}} w_{xy}$, so that the total original and new successes are $\langle w, Hw\rangle$ and $\langle w, H^3 w\rangle$; for $s \ge \frac{1}{2}$, $(h-s)^2(h+2s) \ge 0$ on $[-1, 1]$ gives $\langle w, H^3 w\rangle \ge 3s^2\langle w, Hw\rangle - 2s^3\|w\|^2 \ge s^3$. The new effects $S(P_x^a)^T$ are positive and sum to the identity on the support; unused support is assigned to one legal answer.

Resources. The resource statements follow from Theorem 8.1: a near-winning strategy of loss ε yields a strategy of loss at most 3ε whose two directional rates are equal and below either original rate. The same precomposition argument applies to a fixed-block one-shot Bell-extraction fidelity and to one-way secret key, since each is monotone under deterministic local preprocessing, invariant under full conjugation of protocols, and reversed by the swap. A conjunction of block-agreement predicates is block agreement with tuple labels, which gives the statement for parallel products. □

The inequality $s' \ge s^3$ is the three-link analogue of the Barnum–Knill bound $P_{\mathrm{PGM}} \ge P_{\mathrm{opt}}^2$ for the pretty-good measurement [9]. That bound is the classical-fidelity case of their main theorem, whose proof for ensembles of more than one state was corrected in 2026 (arXiv version 2 of [9]); the proof of $s' \ge s^3$ above is self-contained and does not use it. Squaring a strategy through such correlation maps has ancestry in the analysis of projection games by Dinur, Steurer and Vidick [52, Lemma 6]. The output is generally mixed and is not a maximally entangled or tracial strategy; this is a resource normal form, not a strategy rounding of the kind studied by Paddock and Culf [112, 45]. The theorem supersedes, in its guarantee, an earlier sequential decoder with a game-dependent loss factor; on noisy mixed magic-square strategies that decoder in fact did better (loss ratio about two, against about three for the construction here), so the factor three is a uniform guarantee, not the best decoder on every instance. There is also a dual, Bob-local construction, with Bob effects $T(Q_y^b)^T$, which produces the identical full correlation table and can keep the smaller local dimension; it changes Bob's tuple, so it does not inherit the literal preservation of Bob's source budgets.

8.3 Sharpness and the exact predicate boundary

Theorem 8.3 (Sharpness of the factor three)**.** *The factor three is sharp for deterministic Alice-local channels producing* $\mathrm{Swap}\,\sigma\,\mathrm{Swap} = \bar{\sigma}$ *on* $\bar{B} : B$*, complex conjugation being taken in Bob's computational basis, with the original Bob and his measurements fixed, even allowing an alternative symmetrizing channel and optimal legal decoding. For* $N \ge 2$ *and* $0 < p < 1$ *take* $\rho = N^{-1}\sum_i |i\rangle\langle i| \otimes \beta_i$ *with* $\beta_i = p|0\rangle\langle 0| + (1-p)|i\rangle\langle i|$*; Alice's answers are the* N *one-hot words, Bob is asked a uniform* j *and tests* $|j\rangle\langle j|$*. The original loss is* p/N*, and the best new loss under this contract is*

$$\frac{p}{N}\Big[1 + 2(1-p)\Big(1 - \frac{1}{N}\Big)\Big],$$

whose ratio to p/N *tends to three. At* $N = 3$*,* $p = \frac{1}{8}$ *the loss goes from* $\frac{1}{24}$ *to* $\frac{13}{144}$*, ratio* $\frac{13}{6}$*. Difunctionality is the exact boundary for this channel-and-whole-answer guarantee: a*

nondifunctional relation has a classical source with three accepted corners, of original loss zero, for which the new rejected probability is $\frac{1}{12}$.

Proof. Alice returns her label i; Bob sees 0 with probability p, in which case the two disagree exactly at $j = i$, so the original loss is p/N. Since $\rho_A = I/N$, the channel sends $|i\rangle\langle i|$ to β_i and $\tau = N^{-1}\sum_i \beta_i \otimes \beta_i$, which is classical. On it any POVM of Alice reduces, for this score, to a classical decision on her label. If she sees $i \geq 1$, answering e_i is optimal and loses p/N on that branch. If she sees 0 (probability p), all one-hot answers tie and an answer e_k disagrees with Bob at $j = k$ if he sees 0, and at two queries if he sees $i \neq k$; the conditional loss is $[p + 2(1-p)(1-1/N)]/N$. Adding the branches gives the display.

Now let Λ be any deterministic Alice-local channel with output of Bob's dimension and $\sigma = (\Lambda \otimes \mathrm{id})\rho$ satisfying the symmetry, and write $\alpha_i = \Lambda(|i\rangle\langle i|)$, so $\sigma = N^{-1}\sum_i \alpha_i \otimes \beta_i$. Bob's side is diagonal, so the symmetry makes σ diagonal on Alice's side too, and each α_j, the Bob block $j \geq 1$ divided by $(1-p)/N$, is diagonal. Equality of Alice's marginal with Bob's conjugate marginal gives $\sum_j \alpha_j(i) = 1 - p$ for $i \geq 1$, and the symmetry of the entries $(0, i)$ and $(i, 0)$ gives $(1-p)\alpha_i(0) = p\sum_j \alpha_j(i)$, so $\alpha_i(0) = p$. Hence $\alpha_i(j) = (1-p)W_{ji}$ for a symmetric doubly stochastic W, and $\sigma = (\Lambda_W \otimes \mathrm{id})(\tau)$, where Λ_W fixes 0 and applies W to the other labels: every permitted output is a local degradation of τ, and no legal decoder on it beats the optimal decoder on τ.

For the boundary, take answers a, a' and b, b' with $(a, b), (a', b), (a', b')$ accepted and (a, b') rejected, and the classical state with joint probabilities $P(a, b) = P(a', b) = P(a', b') = \frac{1}{3}$, $P(a, b') = 0$. The canonical output has table $\tau(b, b) = \frac{1}{2}$, $\tau(b, b') = \tau(b', b) = \tau(b', b') = \frac{1}{6}$, and Alice's new effect for a is $\mathrm{diag}(\frac{1}{2}, 0)$, so the rejected pair has probability $\frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$. □

The sharpness does not depend on the choice of conjugation: if $J = UK$ is any antiunitary on Bob's space with $U\bar{U} = \pm I$, applied to both factors, and one requires $\mathrm{Swap}\,\sigma\,\mathrm{Swap} = (U \otimes U)\bar{\sigma}(U \otimes U)^*$, then every α_i is diagonal in the rotated basis $\{U|m\rangle\}$ and rotating Alice's output by U^* reduces to the case above. The example is separable and its game has a perfect classical strategy; its force is a sharp local normal form, not an entanglement bound. Sharpness is not claimed for an unrestricted rescaling of demand profiles, for operations that change Bob, or for two-party preprocessing. With an illegal all-zero answer allowed, the ratio in the example would drop to $2(1-p)$; the original legal alphabet is essential.

8.4 Synchronous games, and what the directions share

Synchronous predicates need not be difunctional off the diagonal, and the factor three need not apply to them. The tested diagonals still give a fixed-game comparison.

Proposition 8.4 (Synchronous repair)**.** *Let G be a game in class (S). There is C_G such that, for every finite-dimensional strategy of loss ε, the construction of Theorem 8.2 gives a strategy of loss at most $\varepsilon + C_G\sqrt{\varepsilon}$ with equal directional rates below either original rate. Hence $d_G^{\leftarrow}(\varepsilon + C_G\sqrt{\varepsilon}) \leq d_G^{\rightarrow}(\varepsilon)$ and symmetrically.*

Proof. For a diagonal question x with success s_x, put $d_{xa} = \langle P_{xa}, (I - TS)P_{xa}\rangle_{\rho_A}$ and $d_x = \sum_a d_{xa}$. By Cauchy–Schwarz, $\sum_a \|SP_{xa}\|^2 \geq (\sum_a \langle SP_{xa}, Q_{xa}\rangle)^2 / \sum_a \|Q_{xa}\|^2 \geq s_x^2$, so $d_x \leq 1 - s_x^2 \leq 2(1 - s_x) \leq 2\varepsilon/\mu(x, x)$. The new probability of (a, b) at (x, y) differs from the old by $\langle (TS - I)P_{xa}, T(Q_{yb})\rangle$, which, since $(I - TS)^2 \leq I - TS$, is at most $\sqrt{d_{xa}}\,\|T(Q_{yb})\|$ in modulus. Summing over the finitely many weighted effects with Cauchy–Schwarz bounds the total change of loss by $C_G\sqrt{\varepsilon}$. □

Corollary 8.5 (Directional agreement of the spectra)**.** *For the agreement games of Theorems 11.9, 12.4 and 13.7, the reverse infima have the same asymptotic exponents and leading*

logarithmic laws as the forward ones, by Theorem 8.2 with a universal rescaling of the loss. For the exact sources of Section 14, Theorems 6.10 and 8.1 and the separate pure upper constructions give the stated orders in both physical directions. For an arbitrary game in class (S), Proposition 8.4 transfers a logarithmic floor between the directions but does not supply the sharp constant. None of this asserts that the two rates of a given original state are equal.

The native class laws of Section 13.5 keep their conditions, constants and classes under the applicable closure (Section 1.6).

III. The spectrum

9 Stable spectral transport and the Slofstra–Vidick game

Slofstra and Vidick built a linear system game G_{SV} with $C\varepsilon^{-1/6} \le \mathcal{E}(G_{\mathrm{SV}},\varepsilon) \le C'\varepsilon^{-1/2}$ from the group [134, Proposition 3.2]

$$K = \langle a,b,c,x,y \;\mid\; xyx^{-1}y^{-2},\; xcx^{-1}c^{-1},\; yay^{-1}b^{-1},\; yby^{-1}a^{-1},\; c^{-1}ab,\; a^2,\; b^2,\; c^2\rangle, \quad (9.1)$$

in which c is a central involution, and whose profile satisfies $C\varepsilon^{-2/3} \le \mathrm{hlp}(c;\delta,\varepsilon) \le C'/\varepsilon$. Our companion paper shows that both lower exponents are the upper ones: $\mathrm{hlp}(c;\delta,\varepsilon) = \Theta(1/\varepsilon)$ and $\mathcal{E}(G_{\mathrm{SV}},\varepsilon) = \Theta(\varepsilon^{-1/2})$ [101, Theorem F]. We recall that argument and add the entropy: the least entanglement entropy, and the least entanglement of formation, are $\frac{1}{2}\log(1/\varepsilon) + O(1)$ (Theorem 9.10).

The picture is as follows. Let μ be the spectral measure of $\varphi(y)$ in the state. The relation $xyx^{-1} = y^2$ makes μ nearly invariant under the doubling map $\vartheta \mapsto 2\vartheta$ of the circle, and, where c acts as -1, the relations $yay^{-1} = b$ and $c = ab$ make it nearly invariant under the half-turn $\vartheta \mapsto \vartheta + \pi$. Averaging over the two inverse branches of the doubling map contracts the Wasserstein distance by a factor 2, so a measure that is nearly invariant under both maps is close to Haar measure, while a measure with d atoms is at distance at least $\pi/(2d)$ from it. All quantities are measured in one state; the Wasserstein-1 distance, rather than the quadratic one, is what makes a change of state cost only a linear error.

9.1 Measures on the circle

Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ with the geodesic distance $d(\vartheta,\vartheta') \in [0,\pi]$, and let h be the normalized Haar measure. For probability measures μ,ν on $\mathbb{T}$ put

$$W_1(\mu,\nu) = \sup\Big\{ \int f\,d\mu - \int f\,d\nu :\; f:\mathbb{T}\to\mathbb{R} \text{ is 1-Lipschitz}\Big\}.$$

By Kantorovich–Rubinstein duality [146, Chapter 5] this is the least mean distance of a coupling of μ and ν, and the least is attained. Until Section 9.3 we use only the triangle inequality and the elementary bound $W_1(\mu,\nu) \le \int d\,d\gamma$ for every coupling γ of μ and ν, which holds because $\int f\,d\mu - \int f\,d\nu = \int (f(\vartheta) - f(\vartheta'))\,d\gamma$. Let $D(\vartheta) = 2\vartheta$ and $S(\vartheta) = \vartheta + \pi$, with push-forwards D_*, S_*.

The two square roots of $e^{i\vartheta}$ have angles $\vartheta/2$ and $\vartheta/2 + \pi$, a set that does not depend on the representative of ϑ. Let P replace each point by the uniform measure on its two square roots: $\int f\,d(\mathsf{P}\mu) = \int \mathsf{P}^* f\,d\mu$ with $\mathsf{P}^* f(\vartheta) = \frac{1}{2}(f(\vartheta/2) + f(\vartheta/2+\pi))$.

Lemma 9.1. *For probability measures μ,ν on $\mathbb{T}$:*

(a) $\mathsf{P}h = h$ *and* $\mathsf{P}(D_*\mu) = \frac{1}{2}(\mu + S_*\mu)$*;*

(b) $W_1(\mathsf{P}\mu, \mathsf{P}\nu) \le \frac{1}{2}W_1(\mu,\nu)$*;*

(c) $W_1(\mu, h) \le W_1(D_*\mu, \mu) + W_1(S_*\mu, \mu)$*;*

(d) if μ *has at most* d *atoms, then* $W_1(\mu, h) \ge \pi/(2d)$*.*

This is [101, Lemma 11.7]; part (c) follows from (a), (b) and the triangle inequality, since P is a contraction for W_1 fixing h, and part (d) from the 1-Lipschitz function $d(\cdot, \operatorname{supp}\mu)$.

The operator P is the transfer operator of the doubling map, and (b) is an averaged form of the square-root lifting that Slofstra and Vidick use in operator norm [134, Claims 3.6 and 3.7]. Part (d) needs no lower bound on the masses of the atoms. More generally, if a Markov operator contracts W_1 by $\lambda < 1$ and fixes h, then $W_1(\mu, h) \le W_1(\mu, \mathsf{P}\mu)/(1-\lambda)$; we need only the case above.

For a unitary Y and a density matrix ρ on $\mathbb{C}^d$ let μ_Y^ρ be the spectral measure of Y in ρ, the probability measure on $\mathbb{T}$ with $\int f\, d\mu_Y^\rho = \operatorname{Tr}(\rho f(Y))$, where $f(Y)$ applies f to the angles of the eigenvalues. It has at most d atoms, and

$$\mu_{Y^2}^\rho = D_*\mu_Y^\rho, \qquad \mu_{-Y}^\rho = S_*\mu_Y^\rho, \qquad \mu_{UYU^*}^\rho = \mu_Y^{U^*\rho U} \quad (U \text{ unitary}). \tag{9.2}$$

Lemma 9.2. *Let* P, Q, Y, U *be unitaries on* $\mathbb{C}^d$ *and* $\xi = \rho^{1/2}$*.*

(a) There is a coupling ν *of* μ_P^ρ *and* μ_Q^ρ*, viewed on the unit circle, with* $\int |z - z'|^2\, d\nu(z,z') = \|P\xi - \xi Q\|_2^2$*. Hence* $W_1(\mu_P^\rho, \mu_Q^\rho) \le \frac{\pi}{2}\|P\xi - \xi Q\|_2$*.*

(b) $W_1(\mu_Y^{U^*\rho U}, \mu_Y^\rho) \le \pi\|[U,\xi]\|_2$*.*

This is [101, Lemma 11.8]. The coupling in (a) is $\nu = \sum_{i,j}\|E_i\xi F_j\|_2^2\delta_{(z_i, z'_j)}$ for spectral decompositions $P = \sum_i z_i E_i$ and $Q = \sum_j z'_j F_j$, and (b) follows from Lemma 2.6.

The coupling in (a) is Connes' joint distribution of left and right multiplication [41, Lemma 1.2.6], as used in [134, (5.4)]; for $\rho = I/d$ it gives the Hoffman–Wielandt inequality. Part (b) is the reason for working with W_1: for $\rho = \operatorname{diag}(\frac{1}{2}+\delta, \frac{1}{2}-\delta)$, $U = \sigma_X$ and $Y = \operatorname{diag}(1,-1)$, the state commutator is $\|[U,\xi]\|_2 = \sqrt{2}(\sqrt{1/2+\delta} - \sqrt{1/2-\delta}) = 2\delta + O(\delta^3)$, while the chordal W_2 distance between the two spectral measures is $\sqrt{8\delta}$.

9.2 The profile of K

Theorem 9.3 ([101, Theorem 11.9 and Proposition 11.10])**.** *Let* φ *assign unitaries* A, B, C, X, Y *on* $\mathbb{C}^d$ *to* a, b, c, x, y*, and let* ρ *be a density matrix on* $\mathbb{C}^d$ *with* $\xi = \rho^{1/2}$*. Suppose that, for some* $e > 0$*,*

(i) every relator r *in (9.1) has* $\|(\varphi(r) - I)\xi\|_2 \le e$*;*

(ii) $\|[\varphi(s), \xi]\|_2 \le e$ *for* $s \in \{a, b, c, x, y\}$*;*

(iii) $\|(C + I)\xi\|_2 \le e$*.*

Then $W_1(\mu_Y^\rho, h) \le \frac{29}{2}\pi e$*, and hence* $d \ge 1/(29e)$*. In the notation of Definition 4.2,* $\operatorname{hlp}_K^\rho(c; e) \ge 1/(29e)$*.*

The bound on W_1 is [101, Proposition 11.10]: the relation $xyx^{-1} = y^2$ gives $W_1(D_*\mu, \mu) \le 5\pi e$ through a coupling in the state and a change of state by X, the relations $yay^{-1} = b$, $c = ab$ and $C \approx -I$ give $W_1(S_*\mu, \mu) \le \frac{19}{2}\pi e$ in the same way, and Lemma 9.1(c) adds the two. Then Lemma 9.1(d) gives the dimension bound [101, Theorem 11.9]. Together with an explicit flat model [101, Example 11.11], a variant of that of [134, Lemma 3.8] of dimension $2\lceil \pi/(\sqrt{2}e)\rceil$, this pins down the profile of c.

Corollary 9.4 (The profile of c; [101, Corollary 11.12])**.** *For every $e > 0$,*

$$\frac{1}{29e} \;\le\; \mathrm{hlp}_K^{\rho}(c;e) \;\le\; \mathrm{hlp}_K(c;2,e) \;\le\; 2\Big\lceil \frac{\pi}{\sqrt{2}\,e} \Big\rceil .$$

For each $0 < \delta \le 2$ there are constants $c_\delta, C_\delta > 0$ with $c_\delta/\varepsilon \le \mathrm{hlp}_K(c;\delta,\varepsilon) \le C_\delta/\varepsilon$ for $0 < \varepsilon \le 2$.

This improves the lower bound $C\varepsilon^{-2/3}$ of [134, Lemma 3.5 and Proposition 3.2] to the order of their upper bound; their argument passes from the normalized norm to the operator norm, which costs a factor $\sqrt{d}$.

9.3 Entropy from transport

Theorem 9.3 counts the atoms of a spectral measure. Atoms may carry very unequal weights, and a spectral measure close to Haar measure can belong to a pure state (Example 9.7). What forces entropy is the state commutator of Y: it makes the broad spectral law of Y readable on both halves of a purification, with nearly equal outcomes, and the correlation between the two readouts is paid for by entanglement.

For $m \ge 2$ cut $\mathbb{T}$ into the arcs $I_k = [2\pi k/m, 2\pi(k+1)/m)$, $k \in \mathbb{Z}_m$, and let $\mathrm{bin}(\vartheta) = k$ for $\vartheta \in I_k$. For $k \in \mathbb{Z}_m$ let $|k|_m$ be its distance to 0 in the cycle $\mathbb{Z}_m$. If $d(\vartheta,\vartheta') = t$, the shorter arc between them meets at most $tm/(2\pi)+2$ of the arcs I_k, so $|\mathrm{bin}(\vartheta) - \mathrm{bin}(\vartheta')|_m \le 1 + tm/(2\pi)$. For $a > 0$ and $t \ge 0$ put $G_a(t) = \log\coth(a/2) + at/\ln 2$.

Lemma 9.5. *If D is a random variable in $\mathbb{Z}_m$ with $\mathbb{E}|D|_m \le t$, then $H(D) \le G_a(t)$ for every $a > 0$.*

Proof. Represent D by an integer $\tilde{D}$ with $|\tilde{D}| = |D|_m$; this is injective, so $H(\tilde{D}) = H(D)$. The weights $q_j = e^{-a|j|}/\coth(a/2)$, $j \in \mathbb{Z}$, form a probability distribution, and by Gibbs' inequality $H(\tilde{D}) \le -\sum_j \Pr[\tilde{D}=j]\log q_j = \log\coth(a/2) + \frac{a}{\ln 2}\mathbb{E}|\tilde{D}|$. □

A probability measure ν on $\mathbb{T}$ is (κ,s)-*spread*, for $\kappa \ge 1$ and $0 < s \le 1$, if $\nu(I) \le (\kappa|I|/2\pi)^s$ for every arc I of length $|I|$. Haar measure is $(1,1)$-spread.

Theorem 9.6 (Entropy from transport)**.** *Let Y be a unitary and ρ a density matrix on $\mathbb{C}^d$, $\xi = \rho^{1/2}$, $\mu = \mu_Y^{\rho}$, and suppose $W_1(\mu,\nu) \le \delta$ for a (κ,s)-spread ν, and $\|[Y,\xi]\|_2 = c$. For every integer $m \ge 2$ and all $a, b > 0$,*

$$S(\rho) \;\ge\; s\log m - s\log\kappa - G_a\Big(1 + \frac{m\delta}{2\pi}\Big) - G_b\Big(1 + \frac{mc}{4}\Big).$$

In particular $S(\rho) \ge s\log(1/(\delta+c)) - s\log\kappa - 8$; for $\nu = h$, $S(\rho) \ge \log(1/(\delta+c)) - 8$.

Proof. Let $Y = \sum_z z E_z$ be the spectral decomposition and $P_k = \sum_{\arg z \in I_k} E_z$. On the canonical purification ψ_ρ (Section 4.1) let Bob measure $\{P_k\}$, with outcome B, and Alice measure $\{P_k^T\}$, with outcome A. Through the isometry Θ, $\Pr[A=k, B=j] = \|(P_k^T \otimes P_j)\psi_\rho\|^2 = \|P_j \xi P_k\|_2^2$, and B has law $\Pr[B=j] = \mathrm{Tr}(P_j\rho) = \mu(I_j)$.

Holevo. Alice's measurement leaves Bob with an ensemble of states whose average is ρ. By Holevo's bound [71], the mutual information between A and the outcome of any measurement by Bob is at most $S(\rho)$ minus the average entropy of the ensemble, so $I(A:B) \le S(\rho)$.

The two readouts nearly agree. Measuring the finer projections E_w^T and E_z instead gives the law $\|E_z \xi E_w\|_2^2$, of which (A,B) is the binning, and by Lemma 9.2(a) with $P = Q = Y$, $\sum_{z,w} |z-w|^2 \|E_z\xi E_w\|_2^2 = \|Y\xi - \xi Y\|_2^2 = c^2$. The arc length is at most $\frac{\pi}{2}|z-w|$, so its mean is at most $\pi c/2$, and $\mathbb{E}|B-A|_m \le 1 + mc/4$. Since B is determined by A and $B-A$, Lemma 9.5 gives $H(B \mid A) \le H(B-A) \le G_b(1+mc/4)$.

Bob's readout is spread. By Kantorovich–Rubinstein duality [146, Chapter 5] there is a coupling of ν and μ with mean distance at most δ. Let (ϑ, ϑ') have this law, $U = \mathrm{bin}(\vartheta)$ and $V = \mathrm{bin}(\vartheta')$. Every arc I_k has $\nu(I_k) \le (\kappa/m)^s$, so $H(U) \ge s\log m - s\log\kappa$; V has the law of B; and $\mathbb{E}|U - V|_m \le 1 + m\delta/(2\pi)$, so $H(B) = H(V) \ge H(U) - H(U \mid V) \ge s\log m - s\log\kappa - G_a(1 + m\delta/(2\pi))$.

Hence $S(\rho) \ge I(A:B) = H(B) - H(B \mid A)$, which is the first claim. For the second let $\alpha = \delta + c$. If $\alpha > \frac{1}{2}$ the claim is trivial. Otherwise take $m = \lfloor 1/\alpha \rfloor \ge 2$, so $s\log m \ge s\log(1/\alpha) - 1$, and $a = b = \frac{1}{2}$. Then $1 + m\delta/(2\pi) \le \frac{5}{4}$ and $1 + mc/4 \le \frac{5}{4}$. Using $\coth x \le 1 + 1/x$, $\log 5 < \frac{7}{3}$ and $\ln 2 > \frac{2}{3}$, $G_{1/2}(\frac{5}{4}) < \frac{7}{3} + \frac{15}{16} = \frac{157}{48}$, and the total loss is below $1 + 2 \cdot \frac{157}{48} < 8$. □

Example 9.7 (Spread alone gives no entropy)**.** Let $Y = \mathrm{diag}(e^{2\pi i j/d})_{0 \le j < d}$ and $\rho = |\psi\rangle\langle\psi|$ with $\psi = d^{-1/2}\sum_j e_j$. The spectral measure is uniform on the d-th roots of unity, so $W_1(\mu, h) = \pi/(2d)$ (moving each point of the circle to the nearest root couples h to μ at this cost, and Lemma 9.1(d) gives the reverse bound). Yet $S(\rho) = 0$. Here $\|[Y, \xi]\|_2^2 = 2 - 2|\langle\psi, Y\psi\rangle|^2 = 2$, so the state commutator in Theorem 9.6 cannot be dropped.

Corollary 9.8 (The entropy profile of K)**.** *Under the hypotheses of Theorem 9.3, if* $0 < e \le \frac{1}{2}$ *then* $S(\rho) \ge \log(1/e) - 10$. *Consequently the infimum* $h_K^S(e)$ *of* $S(\rho)$ *over all such models satisfies*

$$\log(1/e) - 10 \;\le\; h_K^S(e) \;\le\; \log(1/e) + 3 \qquad (0 < e \le \tfrac{1}{2}).$$

Proof. By Theorem 9.3, $W_1(\mu_Y^\rho, h) \le \frac{29}{2}\pi e$, and $c = \|[Y, \xi]\|_2 \le e$. Apply Theorem 9.6 with $m = \lfloor 1/e \rfloor \ge 2$, $a = \frac{1}{9}$ and $b = \frac{1}{2}$. Then $1 + m\delta/(2\pi) \le \frac{33}{4}$ and $1 + mc/4 \le \frac{5}{4}$. Since $\coth(1/18) \le 19$, $19^4 < 2^{17}$ and $\ln 2 > \frac{2}{3}$, $G_{1/9}(\frac{33}{4}) < \frac{17}{4} + \frac{11}{8}$; as before $G_{1/2}(\frac{5}{4}) < \frac{7}{3} + \frac{15}{16}$; and $\log m \ge \log(1/e) - 1$. The total loss is below $1 + \frac{17}{4} + \frac{11}{8} + \frac{7}{3} + \frac{15}{16} = \frac{475}{48} < 10$. For the upper bound, the flat model of [101, Example 11.11] with $m = \lceil \pi/(\sqrt{2}e) \rceil$ and the flat state has $S = \log(2m) \le \log((\sqrt{2}\pi + 1)/e) < \log(1/e) + 3$. □

9.4 The game

Let $G_{\mathrm{SV}} = G(A, b)$ be the linear system game of [134, Theorem 1.1], with the question distribution of [134] (Definition 4.1), m equations, n variables and solution group Γ. Its construction [134, Propositions 6.1 and 6.2] supplies a homomorphism $\Psi : F(a, b, c, x, y) \to F(x_1, \dots, x_n, J)$ that descends to an embedding of K into Γ, and a word z in $x_1, \dots, x_n, J$, the image of a generator t with $t^2 = e$ and $tct = cJ$, such that $[\Psi(c), z] = J$ in Γ [134, proof of Proposition 6.2]; so z represents an involution. The proof there prints the second relation, for its central element w, as $twt = J$. This is a misprint for $twt = wJ$: the representation used there, with t exchanging the two blocks, $w \mapsto -I \oplus I$ and $J \mapsto -I$, satisfies $twt = wJ$, which is the relation $tat = Ja$ of [132, Lemma 5.2] that the proof invokes, whereas $twt = J$ would force $w = J$.

The upper bound of [134, Theorem 1.1] rests on the lift in the same proof, which needs a $2d$-dimensional representation of the intermediate group with $c \mapsto (-I_d) \oplus I_d$. The value $\omega_q(G_{\mathrm{SV}}) = 1$ does not: [134] obtain it from the non-triviality of J in approximate representations of Γ, through [132, Lemma 5.2 and Proposition 4.2] and [134, Corollary 5.3], and the tensor-power argument of [132, Lemma 5.2] does not need a balanced marked involution. The group K has the two non-involutive generators $y < x$, and the literal lifts of [132, Proposition 4.8] give $(-I_d) \oplus I_{3d}$ instead, for which the relation $tct = cJ$ has no approximate solution (see before Proposition 5.10). A balanced lift exists for the flat models of [101, Example 11.11], and our companion paper proves it [101, Lemma 11.13]: for these models of dimension d, the solution group Γ has flat models with $J \mapsto -I$ of dimension $8d$ at precision $O(1/d)$. The idea is that X is a permutation matrix there, so once y has been eliminated as in [132, Proposition 4.8], x can be eliminated without doubling the dimension,

as the product of the involutions $\tau \oplus \tau$ and $\tau X \oplus \tau X$, where the involutive permutation τ reverses each cycle of X; the marked involution then stays balanced. These models give the upper bounds of Theorem 9.9 and Theorem 9.10(c), so [134, Theorem 1.1] holds as stated, with only the lift in its proof replaced. The script `verify/check_coefficients.py` replays the reversal, the balance of c and the lifted residuals, independently of the companion's own check.

Theorem 9.9 ([101, Theorem 11.14])**.** *There are constants $c_0, C_0 > 0$ such that $c_0\,\varepsilon^{-1/2} \le \mathcal{E}(G_{\mathrm{SV}}, \varepsilon) \le C_0\,\varepsilon^{-1/2}$ for $0 < \varepsilon \le 1$.*

The upper bound is that of [134, Theorem 1.1], obtained through the lift above. The lower bound [101, Section 11.2] combines Theorem 4.3, the word z, which keeps $\varphi(\Psi(c))$ visibly away from I in Bob's state, the passage to the sector where its rounding is -1 (Lemma 2.10), and Theorem 9.3. It bounds the Schmidt rank, and so the local dimension: $\log \mathcal{E}(G_{\mathrm{SV}}, \varepsilon) = \frac{1}{2}\log(1/\varepsilon) + O(1)$. It is not a bound on the entanglement entropy, since Theorem 9.3 counts atoms of a spectral measure, whose masses may be unequal. The entropy is the subject of the next theorem.

Theorem 9.10 (Entanglement entropy and formation for G_{SV})**.** *There are constants $C, \varepsilon_0 > 0$ such that:*

(a) every pure-state strategy for G_{SV} that wins with probability at least $1-\varepsilon$, $0 < \varepsilon \le \varepsilon_0$, has $S(\rho_B) \ge \frac{1}{2}\log(1/\varepsilon) - C$;

(b) for every state σ on $H_A \otimes H_B$ and measurements that win with probability at least $1-\varepsilon$, $0 < \varepsilon \le 1$, $E_{\mathrm{F}}(\sigma) \ge \frac{1}{2}\log(1/\varepsilon) - C$;

(c) for $0 < \varepsilon \le 1$ there are pure-state strategies that win with probability at least $1-\varepsilon$ with entanglement entropy at most $\frac{1}{2}\log(1/\varepsilon) + C$.

So the least entanglement entropy of a pure-state strategy, and the least entanglement of formation of any strategy, are both $\frac{1}{2}\log(1/\varepsilon) + O(1)$.

Proof. (a) Take a strategy that wins with probability at least $1-\varepsilon$, and let φ and ρ on $\mathbb{C}^r$ be as in Theorem 4.3: ρ is Bob's reduced state on its support, and (φ, ρ) is a model of Γ at precision $e_0 = \max(s,4)\theta$, $\theta = 2\sqrt{mn\varepsilon}$, with $\varphi(J) = -I$. Apply Proposition 5.5 with K, presented by (9.1) and with distinguished involution c, in place of G, with Γ in place of H, $\psi = \Psi$ and the word z above; c is a central involution of K, Ψ induces a homomorphism $K \to \Gamma$, and z represents an involution with $[\Psi(c), z] = J$. This gives unitaries on $\mathbb{C}^r$ that satisfy (i)–(iii) of Theorem 9.3, with the same ρ, at precision $e = C_K e_0$ for a constant C_K. If $e \le \frac{1}{2}$, Corollary 9.8 gives

$$S(\rho_B) = S(\rho) \ge \log(1/e) - 10 = \tfrac{1}{2}\log(1/\varepsilon) - \log\bigl(2\max(s,4)C_K\sqrt{mn}\bigr) - 10.$$

(b) Let $f(x) = \frac{1}{2}\log(1/x) - C$, enlarging C so that $C \ge \frac{1}{2}\log(1/\varepsilon_0)$; then (a) and $S \ge 0$ give $S(\rho_B) \ge f(\ell)$ for every pure strategy with losing probability $\ell \in (0,1]$. No pure strategy has $\ell = 0$, by Theorem 9.9. Let $\sigma = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ be any decomposition into pure states. The losing probability is linear in the state, so with the same measurements the components have losing probabilities $\ell_j > 0$ with $\sum_j p_j \ell_j \le \varepsilon$. Since f is convex and decreasing, $\sum_j p_j S((\psi_j)_B) \ge \sum_j p_j f(\ell_j) \ge f(\sum_j p_j \ell_j) \ge f(\varepsilon)$. Taking the infimum over decompositions proves (b).

(c) By the upper bound of Theorem 9.9 there are strategies of local dimension at most $C_0\varepsilon^{-1/2}$, and the entanglement entropy is at most the logarithm of the local dimension. □

The same model also gives $r \geq 1/(29 C_K e_0)$ by Theorem 9.3, a second proof of the lower bound of Theorem 9.9 that keeps the whole state instead of passing to the sector where c acts as -1; the sector would cost a factor $p \geq \frac{1}{16}$ in front of the entropy, which is why the corner map is used here. The two results are different statements: Theorem 9.9 bounds the local dimension, $\Theta(\varepsilon^{-1/2})$, and Theorem 9.10 the entropy, $\frac{1}{2}\log(1/\varepsilon) + O(1)$; neither bounds usable entanglement (Section 1.6).

10 Forced branching: the lamp sources

This section and the next build source groups with prescribed polynomial profiles. The mechanism is branching forced at each magnification. A relation $xyx^{-1} = y^m$ makes the spectral law of y nearly invariant under multiplication of the angle by m, and a finite set of involutions permuted cyclically by y makes it nearly invariant under a rotation by $1/q$ of a turn. A law invariant under both maps puts mass at most q^{-n} on an arc of length $q^{-1}m^{-n+1}$, so it satisfies a Frostman bound of exponent $s = \log_m q$ and has dimension at least s; Haar measure, for instance, is invariant and has dimension one. A spectral law that is nearly invariant is close to such a law, so it needs about e^{-s} atoms and $s\log(1/e)$ bits of entropy.

Turns. In this section and the next we identify $\mathbb{T}$ with $\mathbb{R}/\mathbb{Z}$ through $\vartheta = 2\pi t$ and measure lengths and Wasserstein distances in turns: $d(t,u) \in [0, \frac{1}{2}]$, and W_1 is $1/(2\pi)$ times the distance of Section 9. Lemma 9.2 then reads $W_1(\mu_P^\rho, \mu_Q^\rho) \leq \frac{1}{4}\|P\xi - \xi Q\|_2$ and $W_1(\mu_Y^{U^*\rho U}, \mu_Y^\rho) \leq \frac{1}{2}\|[U,\xi]\|_2$, and a (κ, s)-spread law satisfies $\nu(I) \leq (\kappa|I|)^s$. Put $D(t) = mt$ and $S(t) = t + 1/q$ modulo 1.

10.1 The sources

Definition 10.1 (Lamp sources)**.** For integers $q \geq 1$ and $m \geq 2$, the *lamp source* $L_{q,m}$ has generators $z_0, \dots, z_{q-1}, w, y, x$ and the relators

$$z_j^2, \quad w^2, \quad [z_j, z_k], \quad [z_j, w], \quad z_0 z_1 \cdots z_{q-1} w,$$
$$y z_j y^{-1} z_{j+1}^{-1} \ (j \bmod q), \qquad xwx^{-1}w^{-1}, \qquad xyx^{-1}y^{-m},$$

with distinguished involution w. The row makes w commute with the lamps z_j, and conjugating it by y permutes its factors cyclically, so w is central. Its profiles are those of Example 5.1(a): a model at precision e is a family of unitaries Z_j, W, Y, X and a density matrix ρ, $\xi = \rho^{1/2}$, with every relator of defect at most e in the state, every generator of state commutator at most e, and $\|(W + I)\xi\|_2 \leq e$.

For $q = 2$ the row gives $z_1 = z_0 w$, and $L_{2,2}$ is the group K of Slofstra and Vidick (9.1), with a, b, c the two lamps and w. The power relation is a dilation of the kind of Slofstra's HNN constructions [132, Section 5] and of Fu's self-tests [63].

Theorem 10.2 (One lamp source)**.** *Let $q = 2^r$ with $r \geq 1$, let q divide m, and put $s = \log_m q$. There are constants $c, C > 0$ such that every model of $L_{q,m}$ at precision $e \leq \frac{1}{2}$ on $\mathbb{C}^d$, with density matrix ρ, has*

$$d \ \geq \ \operatorname{rank}\rho \ \geq \ c\, e^{-s} \qquad \text{and} \qquad S(\rho) \ \geq \ s\log(1/e) - C,$$

and for every $e \leq \frac{1}{2}$ there is a model at precision e with the flat state, of dimension at most Ce^{-s} and entropy at most $s\log(1/e) + C$. So $h_{L_{q,m}}(e) = \Theta(e^{-s})$ and $h^S_{L_{q,m}}(e) = s\log(1/e) + O(1)$.

The proof occupies the rest of the section. Its steps are a rotation implemented on the whole state (Section 10.2), an invariant law nearby (Section 10.3), the spread of every invariant law (Section 10.4), and matching models (Section 10.5).

10.2 A rotation on the whole state

Work first in the group algebra of $L_{q,m}$, before $w = -1$ is imposed. For a sign vector $\chi \in \{\pm 1\}^q$ put $P_\chi = 2^{-q}\prod_j (1 + \chi_j z_j)$. These are commuting self-adjoint idempotents with $P_\chi P_{\chi'} = 0$ for $\chi \neq \chi'$, $\sum_\chi P_\chi = 1$ and $wP_\chi = (\prod_j \chi_j)P_\chi$, so

$$\sum_{\chi \in \Omega_-} P_\chi = \tfrac{1}{2}(1 - w), \qquad \Omega_- = \Big\{\chi : \prod_j \chi_j = -1\Big\}.$$

With $(\sigma\chi)_k = \chi_{k-1}$, the conjugacies give $yP_\chi y^{-1} = P_{\sigma\chi}$. Every $\chi \in \Omega_-$ has an orbit of length exactly q: if its least period l is smaller, its first l entries repeat q/l times, an even number when q is a power of two, and the product of all entries is $+1$. Choose a representative χ_O of each orbit $O \subseteq \Omega_-$, put $\zeta = e^{2\pi i/q}$, and

$$T = \sum_O \sum_{a=0}^{q-1} \zeta^a P_{\sigma^a \chi_O}.$$

Lemma 10.3. *In the group algebra, $T^*T = TT^* = \frac{1}{2}(1 - w)$ and $Ty = \zeta yT$. Consequently, for a model of $L_{q,m}$ at precision e there is a unitary U with*

$$\|[U, \xi]\|_2 \le Ce \qquad \text{and} \qquad \|(UYU^* - \zeta Y)\xi\|_2 \le Ce,$$

with C depending only on q and m.

Proof. By orthogonality, $T^*T = TT^* = \sum_O \sum_a P_{\sigma^a\chi_O} = \frac{1}{2}(1 - w)$. Next $yTy^{-1} = \sum_O \sum_a \zeta^a P_{\sigma^{a+1}\chi_O} = \zeta^{-1}T$, since $\sigma^q\chi_O = \chi_O$; so $Ty = \zeta yT$. These identities have finite derivations from the defining relators (expand the products, reduce squares, commute lamps, use the row for w, and the q conjugacies), so in the algebra $L^-_{q,m}$ of Example 5.1(a) the elements $T^*T - 1$, $TT^* - 1$ and $Ty - \zeta yT$ lie in the defining ideal. Let $B = T(\varphi)$. By Lemma 5.2, $\|(B^*B - I)\xi\|_2$, $\|(BB^* - I)\xi\|_2$, $\|(BY - \zeta YB)\xi\|_2$ and $\|[B, \xi]\|_2$ are $O(e)$, and $\|B\|_{\mathrm{op}} = O(1)$. Lemma 2.7 gives a unitary U with $\|(U - B)\xi\|_2$, $\|\xi(U - B)\|_2$ and $\|[U, \xi]\|_2$ of order e. Then $(UY - \zeta YU)\xi = (BY - \zeta YB)\xi + (U - B)\xi Y + (U - B)[Y, \xi] - \zeta Y(U - B)\xi$ is $O(e)$, and $(UYU^* - \zeta Y)\xi = (UY - \zeta YU)U^*\xi$ is $O(e)$ after moving U^* past ξ. □

No sector is selected: the polynomial T combines all sign vectors of odd parity, and the state is never changed. For $q = 2$, $T = P_{(-,+)} - P_{(+,-)}$, which in the negative quotient is the lamp z_0 up to sign, and U is close to it.

Lemma 10.4 (Two approximate invariances)**.** *For a model of $L_{q,m}$ at precision e, the spectral law $\mu = \mu^\rho_Y$ satisfies*

$$\delta_D = W_1(D_*\mu, \mu) \le \tfrac{m+3}{2}\, e, \qquad \delta_S = W_1(S_*\mu, \mu) \le C'e.$$

Proof. Put $Q = XYX^*$ and $R = Y^m$. The relator gives $\|(QR^* - I)\xi\|_2 \le e$, so $\|(Q - R)\xi\|_2 \le e + 2\|[R, \xi]\|_2 \le (2m + 1)e$; with $\|[Q, \xi]\|_2 \le 3e$, $\|R\xi - \xi Q\|_2 \le (2m + 4)e$. By (9.2) and Lemma 9.2 in turns, $W_1(D_*\mu, \mu_Y^{X^*\rho X}) \le \frac{m+2}{2}e$ and $W_1(\mu_Y^{X^*\rho X}, \mu) \le \frac{e}{2}$. For the rotation let U be as in Lemma 10.3: $\mu^\rho_{\zeta Y} = S_*\mu$, $\|\zeta Y\xi - \xi UYU^*\|_2 \le \|(\zeta Y - UYU^*)\xi\|_2 + \|[UYU^*, \xi]\|_2 = O(e)$, and $\mu^\rho_{UYU^*} = \mu_Y^{U^*\rho U}$ is within $\frac{1}{2}\|[U, \xi]\|_2$ of μ. □

10.3 An invariant law nearby

For $m = 2$ and $q = 2$ the invariant law is unique, and Lemma 9.1 gives a direct estimate. In general there are many invariant laws, Cantor-like ones included, and we show instead that a nearly invariant law is close to *some* invariant law, by shadowing a stationary pseudo-orbit.

Lemma 10.5 (Shadowing)**.** *Let q divide m, and let μ be a probability measure on $\mathbb{R}/\mathbb{Z}$ with $\delta_D = W_1(D_*\mu, \mu)$ and $\delta_S = W_1(S_*\mu, \mu)$. There is a probability measure ν with $D_*\nu = \nu$ and $S_*\nu = \nu$ such that*

$$W_1(\mu, \nu) \le b + \frac{\delta_D + b}{m - 1}, \qquad b = \frac{(q-1)\delta_S}{2}.$$

Proof. Symmetrize. Let $\mu_0 = \frac{1}{q}\sum_{k<q} S_*^k\mu$. Since S is an isometry, $W_1(\mu, S_*^k\mu) \le k\delta_S$, so $W_1(\mu, \mu_0) \le b$. As q divides m, $DS = D$, so $D_*\mu_0 = D_*\mu$ and $\delta := W_1(D_*\mu_0, \mu_0) \le \delta_D + b$.

A symmetric coupling. Let η be an optimal coupling of $D_*\mu_0$ and μ_0. Disintegrate μ_0 along the fibres of D; since S preserves every fibre and μ_0, the conditional laws can be replaced by their averages over the powers of S. Choosing (z, v) from η and then u from the conditional law of μ_0 on the fibre $D^{-1}(z)$ gives a law Q of (u, v) with both marginals μ_0, with $Du = z$, so that $\int d(Du, v)\, dQ = \delta$, and invariant under $(u, v) \mapsto (Su, v)$.

A stationary chain. Let $(U_t)_{t\in\mathbb{Z}}$ be the stationary Markov chain with stationary law μ_0 whose consecutive pairs (U_t, U_{t+1}) have law Q. The conditional law of U_t given the whole future $(U_{t+1}, U_{t+2}, \dots)$ is the reverse kernel of Q at U_{t+1}, by the Markov property of the reversed chain; by the symmetry of Q it is invariant under S.

Shadowing. Let $E_t \in [-\frac{1}{2}, \frac{1}{2})$ represent $mU_t - U_{t+1}$ modulo 1, so $\mathbb{E}|E_t| = \delta$, and put

$$V_t = U_t - \sum_{j\ge 0} m^{-j-1} E_{t+j} \pmod 1.$$

The series converges absolutely, and $mV_t - V_{t+1} = mU_t - E_t - U_{t+1} = 0$ modulo 1. The process (V_t) is a fixed function of the shifted process, so it is stationary, and the law ν of V_0 satisfies $D_*\nu = \nu$. Also $\mathbb{E}\, d(U_0, V_0) \le \sum_j m^{-j-1}\mathbb{E}|E_j| = \delta/(m-1)$, so $W_1(\mu_0, \nu) \le \delta/(m-1)$.

Symmetry. E_0 depends on U_0 only through mU_0, which is unchanged when U_0 is replaced by SU_0 because q divides m, and every later E_j depends only on the future. So $V_0 = U_0 - g(mU_0, U_1, U_2, \dots)$ for a measurable g, and replacing U_0 by SU_0 replaces V_0 by SV_0. Conditional on the future, U_0 has an S-invariant law; hence so has V_0, and ν is S-invariant. Finally $W_1(\mu, \nu) \le W_1(\mu, \mu_0) + W_1(\mu_0, \nu)$. □

No uniqueness, ergodicity or continuity of an inverse branch is used, and the pseudo-orbit errors are controlled only on average.

10.4 Every invariant law is spread

Lemma 10.6 (Spread)**.** *Let $q \ge 2$ divide m and $s = \log_m q$. Every probability measure ν on $\mathbb{R}/\mathbb{Z}$ with $D_*\nu = \nu$ and $S_*\nu = \nu$ has no atoms and satisfies $\nu(I) \le (q|I|)^s$ for every arc I; it is (q, s)-spread. Consequently every probability measure λ with at most d atoms has $W_1(\lambda, \nu) \ge d^{-1/s}/(4qm)$.*

Proof. For a point t, the q points $S^k t$ are distinct and all map to Dt, so $q\nu(\{t\}) \le \nu(\{Dt\})$, and iterating gives $\nu(\{t\}) \le q^{-n}$ for every n. For a half-open arc I of length at most $1/q$, the arcs $S^k I$ are disjoint and lie in $D^{-1}(DI)$, so $q\nu(I) \le \nu(D^{-1}(DI)) = \nu(DI)$, where DI is an arc of length $\min(m|I|, 1)$. For $0 < |I| \le 1/q$ let $n = \lfloor \log_m(1/(q|I|)) \rfloor + 1$; the first n images of I have length at most $1/q$ before each step, so $\nu(I) \le q^{-n} < (q|I|)^s$. Longer arcs satisfy the bound trivially, and endpoints do not matter.

Let F be the set of atoms of λ, $r = (2d)^{-1/s}/(2q)$, and $f(t) = d(t, F)$. The points within r of F form at most d arcs of length $2r$, of total ν-mass at most $d(2qr)^s = \frac{1}{2}$. So $\int f\, d\nu \geq r/2$ while $\int f\, d\lambda = 0$, and $W_1(\lambda, \nu) \geq r/2 = 2^{-1/s} d^{-1/s}/(4q) \geq d^{-1/s}/(4qm)$, since $2^{1/s} = m^{\log_q 2} \leq m$. □

Proof of the lower bounds in Theorem 10.2. By Lemmas 10.4 and 10.5 there is an invariant ν with $W_1(\mu, \nu) \leq C''e$. The law μ has at most d atoms, so Lemma 10.6 gives $d \geq (4qmC''e)^{-s}$. In radians ν is (q, s)-spread and $W_1(\mu, \nu) \leq 2\pi C''e$, and $\|[Y, \xi]\|_2 \leq e$, so Theorem 9.6 gives $S(\rho) \geq s \log(1/e) - C$. Finally $\operatorname{rank} \rho \geq 2^{S(\rho)}$. □

The entropy bound uses the state commutator of Y, not only the spread of its spectral law: by Example 9.7 a pure state can have a spread spectral law.

10.5 Matching models

Example 10.7 (Fourier lamps)**.** Let q divide m, $L \geq 1$, $h = m^L - 1$, and index a basis of $\mathbb{C}^{q^L}$ by words $b = (b_0, \ldots, b_{L-1})$ over $\{0, \ldots, q-1\}$. Put

$$t_b = \frac{m}{qh} \sum_{j=0}^{L-1} b_j m^{L-1-j}, \qquad Y = \operatorname{diag}(e^{2\pi i t_b}), \qquad W = -I,$$

let X be the permutation matrix with $X_{b,Rb} = 1$, where R rotates a word one place to the left, and on the first letter let $v_j(a) = q^{-1/2}\zeta^{ja}$, $E_j = |v_j\rangle\langle v_j|$ and $Z_j = (I_q - 2E_j) \otimes I_{q^{L-1}}$. Then all relators of $L_{q,m}$ hold exactly except the lamp conjugacies, whose defects are at most $4\pi(q-1)/(q(m^L - 1))$ in operator norm.

Proof. The E_j are orthogonal rank-one projections with sum I_q, so the Z_j are commuting involutions with $\prod_j Z_j = I - 2\sum_j E_j = -I$, and the row, the squares and the commutators hold; $W = -I$ is central. An integer calculation gives $mt_b - t_{Rb} = (m/q)b_0 \in \mathbb{Z}$, so $XYX^* = Y^m$ exactly. Write $t_b = b_0/q + b_0/(qh) + r_b$, with r_b depending only on $b_1, \ldots, b_{L-1}$, and let $Y_0 = D_0 \otimes \operatorname{diag}(e^{2\pi i r_b})$ with $D_0 = \operatorname{diag}(\zeta^a)$ on the first letter. Then $D_0 E_j D_0^* = E_{j+1}$, so $Y_0 Z_j Y_0^* = Z_{j+1}$, including $j = q-1$, and $\|Y - Y_0\|_{\mathrm{op}} \leq \max_{b_0} |e^{2\pi i b_0/(qh)} - 1| \leq 2\pi(q-1)/(qh)$. Hence $\|Y Z_j Y^* - Z_{j+1}\|_{\mathrm{op}} \leq 2\|Y - Y_0\|_{\mathrm{op}}$, and this is the defect of the relator, since its last factor is unitary. □

Proof of the upper bounds in Theorem 10.2. With the flat state every state commutator vanishes and every defect in the state is at most its operator norm. Choose L least with $4\pi(q-1)/(q(m^L - 1)) \leq e$; then $m^{L-1} \leq C/e$, and the dimension $q^L = q\,(m^{L-1})^s \leq Ce^{-s}$ has entropy $L \log q \leq s \log(1/e) + C$. □

Remark 10.8 (The arithmetic of the lamp count). Write $q = 2^v t$ with t odd. Every odd-parity sign vector has an orbit whose length l satisfies that q/l is odd, so 2^v divides l. Grading each orbit by powers of $e^{2\pi i/2^v}$ gives, by the same proof, a rotation by 2^{-v} of a turn on the whole state; if 2^v divides m, the exponent of $L_{q,m}$ is $\log_m 2^v$, with matching models obtained from 2^v Fourier lamps, each repeated t times. For odd q there is an exact one-dimensional model, with every lamp and w equal to -1 and $x = y = 1$, and the profile is bounded. So the power of two in q is what branches; an odd factor adds nothing. The divisibility of m by the rotation order is used in $DS = D$, in the shadowing and in the models.

10.6 A five-generator source with the same profile

The lamps can be replaced by a single central element of order q. For $q = 2^r$ dividing m, let $G_{q,m} = \langle a, c, w, x, y\rangle$ with the relators a^q, c^q, w^2, the commutators of c with a, x, y, w and of w with a, x, y, the relation $w = c^{q/2}$, and $aya^{-1} = cy$, $xyx^{-1} = y^m$, the marked involution being w. In the group algebra put $P_j = q^{-1}\sum_l \zeta^{-jl}c^l$; these are orthogonal projections commuting with a, x, y, w, with $cP_j = \zeta^j P_j$ and $\sum_{j\ \mathrm{odd}} P_j = \frac{1}{2}(1 - c^{q/2})$. For odd j let k_j be the inverse of j modulo q and put $T = \sum_{j\ \mathrm{odd}} P_j a^{k_j}$. Then $T^*T = TT^* = \frac{1}{2}(1 - w)$ and $Ty = \zeta yT$ in the full group algebra, by finite derivations from the relators, so the conclusion of Lemma 10.3 holds for $G_{q,m}$, and with it every argument of this section. Hence $G_{q,m}$ has minimum ambient dimension and minimum rank $\Theta(e^{-s})$ and minimum entropy $s\log(1/e) + O(1)$, $s = \log_m q$; the abandoned variant with $q = 6$ and $c = -1$ does not create a six-way exponent, since its negative sector can be two-dimensional. We use the lamp sources below because they lie in the restricted linear input class.

10.7 Additivity over copies, without a joint spectral measure

For $i = 1, \dots, k$ fix $q_i = 2^{r_i}$ dividing m_i and a copy of L_{q_i,m_i}, identify the marked involutions, and add to the relators every commutator between generators of different copies. Call this finite presentation $L_{\mathbf{q},\mathbf{m}}$, and put

$$s_i = \log_{m_i} q_i, \qquad s = \sum_i s_i.$$

Theorem 10.9 (Additive profiles)**.** *There are constants $c, C > 0$, depending on the fixed presentation, such that every model of $L_{\mathbf{q},\mathbf{m}}$ at precision $e \le \frac{1}{2}$, with density matrix ρ on $\mathbb{C}^d$, satisfies*

$$S(\rho) \ \ge\ s\log(1/e) - C, \qquad d \ \ge\ \operatorname{rank}\rho \ \ge\ c\,e^{-s},$$

and tensor products of the Fourier-lamp models of Example 10.7 give models of dimension at most Ce^{-s} and entropy at most $s\log(1/e) + C$. So the minimum dimension and the minimum rank are $\Theta(e^{-s})$, and the minimum entropy is $s\log(1/e) + O(1)$.

The difficulty is that the matrices Y_i of different copies commute only approximately, in a weighted norm, and ρ may be singular or far from flat. We do not assign them a joint spectral measure, and we do not round them to commuting matrices. Instead we measure spectral arcs of short words copy by copy, in reverse order of depth, and show that the record has additive entropy and that Alice can predict it at bounded cost.

Notation. Norms are unnormalized Hilbert–Schmidt norms, $\|\xi\|_2 = 1$, and $c(V) = \|[V, \xi]\|_2$. Constants C depend on the fixed sources and on k, never on e, the depth, the dimension or ρ. Spectral angles are in turns.

Lemma 10.10 (Short words)**.** *For each copy there is a unitary U_i with $c(U_i) \le Ce$, $\|(U_iY_iU_i^* - \zeta_iY_i)\xi\|_2 \le Ce$, $\zeta_i = e^{2\pi i/q_i}$, and $\|[U_i, V]\xi\|_2 \le Ce$ for every generator V of another copy. Put $Z_{i,j} = X_i^j Y_i X_i^{-j}$ and $T_{i,j} = X_i^j U_i X_i^{-j}$. For depths $j, l \le N$,*

$$c(Z_{i,j}) + c(T_{i,j}) \le C(N+1)e,$$
$$\|[Z_{i,j}, Z_{h,l}]\xi\|_2 + \|[T_{i,j}, Z_{h,l}]\xi\|_2 \le C(N+1)^3 e \quad (i \ne h),$$
$$\|[Z_{i,j}, Z_{i,l}]\xi\|_2 \le Ce\big(m_i^{|l-j|} + N + 1\big), \qquad \|(T_{i,j}Z_{i,j}T_{i,j}^* - \zeta_i Z_{i,j})\xi\|_2 \le C(N+1)e,$$
$$\|[T_{i,j}, Z_{i,l}]\xi\|_2 \le Ce\big(m_i^{l-j} + N + 1\big) \quad (l > j).$$

Proof. The unitary U_i is that of Lemma 10.3. The selector polynomial B_i uses only copy-i generators, so its commutator with a generator V of another copy has a finite certificate and

is $O(e)$ on ξ; and $[U_i - B_i, V]\xi$ is $O(e)$ because both one-sided differences $(U_i - B_i)\xi$ and $\xi(U_i - B_i)$ are $O(e)$ and $c(V) \le e$.

For unitaries R, S, left multiplication by R and right multiplication by S are commuting unitaries on the Hilbert–Schmidt space, so

$$\|R^a\xi - \xi S^a\|_2 \le a\|R\xi - \xi S\|_2 \qquad (a \ge 0). \tag{10.1}$$

Let $b_d = \|(X^dYX^{-d} - Y^{m^d})\xi\|_2$ for one copy. Then $X^{d+1}YX^{-d-1} - Y^{m^{d+1}} = X(X^dYX^{-d} - Y^{m^d})X^{-1} + ((XYX^{-1})^{m^d} - (Y^m)^{m^d})$. The first term is at most $b_d + 2e$ on ξ. For the second, the relator gives $\|XYX^{-1}\xi - \xi Y^m\|_2 \le (3m+1)e$, so by (10.1) and $c(Y^{mn}) \le mne$ it is at most $(4m+1)m^de$. Hence $b_d \le Cm^de$, with no factor m^{2d}.

Conjugating by $X^{\min(j,l)}$, at cost $O((N+1)e)$, reduces $[Z_j, Z_l]$ to $[Y, X^{|l-j|}YX^{-|l-j|}]$, and the second factor is within $Cm^{|l-j|}e$ of $Y^{m^{|l-j|}}$, which commutes with Y. Applying (10.1) to UYU^* and ζY gives UY^nU^* within $O(ne)$ of ζ^nY^n, and $\zeta^{m^d} = 1$ for $d \ge 1$ because $q \mid m$; comparing with b_d and conjugating by X^j gives the estimate for $[T_j, Z_l]$. The conjugacy estimate is the defining property of U conjugated by X^j. Words of different copies have length $O(N+1)$ and commute after $O((N+1)^2)$ swaps, each with a suffix of length $O(N+1)$. □

For a unitary Z and an offset ϑ uniform modulo $1/q$, let $P_a(Z;\vartheta)$, $a \in \mathbb{Z}_q$, be the spectral projections of Z onto the arcs $[\vartheta + a/q, \vartheta + (a+1)/q)$; multiplying Z by ζ permutes them cyclically, for every offset. For partitions put $C(P) = (\sum_a \|[P_a, \xi]\|_2^2)^{1/2}$, $D(P,V) = (\sum_a \|[P_a, V]\xi\|_2^2)^{1/2}$ and $D(P,Q) = (\sum_{a,b} \|[P_a, Q_b]\xi\|_2^2)^{1/2}$.

Lemma 10.11 (Random spectral cuts)**.** *For unitaries R, S, $\mathbb{E}\sum_a \|P_a(R)\xi - \xi P_a(S)\|_2^2 \le \frac{q}{2}\|R\xi - \xi S\|_2$. Consequently, for unitaries Z, V, B,*

$$\mathbb{E}C(P(Z))^2 \le C_q c(Z), \qquad \mathbb{E}D(P(Z),V)^2 \le C_q\big(\|[Z,V]\xi\|_2 + c(Z) + c(V)\big),$$
$$\mathbb{E}D(P(Z),P(B)) \le C\big(\|[Z,B]\xi\|_2 + c(Z) + c(B)\big)^{1/4},$$

and if V conjugates Z approximately to ζZ, with π the induced cyclic permutation of the arcs, $\mathbb{E}\sum_a \|(P_a(Z)V - VP_{\pi(a)}(Z))\xi\|_2^2 \le C_q(\|(VZV^* - \zeta Z)\xi\|_2 + c(Z) + c(V))$.

Proof. The weights $\|E_R(dt)\xi E_S(du)\|_2^2$ form a coupling of the ρ-spectral laws of R and S with $\int |e^{2\pi it} - e^{2\pi iu}|^2 = \|R\xi - \xi S\|_2^2$ (Lemma 9.2). The left side of the first claim is twice the coupling mass of pairs in different arcs, and a random offset separates t from u with probability at most $q\,d(t,u)$, where $d(t,u) \le \frac{1}{4}|e^{2\pi it} - e^{2\pi iu}|$; Cauchy–Schwarz gives the bound. Take $R = S = Z$ for the first consequence, and $R = VZV^*$, $S = Z$ or $S = \zeta Z$ for the second and the covariant one, converting $P_a(R)\xi - \xi P_a(S)$ into the stated commutators at the cost of $C(P(Z))$ and $c(V)$; all quantities are bounded, so squares may be absorbed into first powers. For the third, the second applied with Z as control bounds $\mathbb{E}\sum_b \|[Q_b, Z]\xi\|_2^2$ for the arcs Q_b of B; applying it again to the arcs of Z with the unitary controls $2Q_b - I$, Cauchy–Schwarz over b and Jensen give the fourth root. □

Lemma 10.12 (Weighted histories)**.** *Let $P_1, \dots, P_M$ be projective partitions with at most Q outcomes, $c_t = C(P_t)$, $c = \sum_t c_t$, $d_{t,s} = D(P_t, P_s)$, and let K_h be the Kraus products of the sequential measurement, in physical order. Then $\sum_h K_h^*K_h = \sum_h K_hK_h^* = I$ and, for every operator F,*

$$\Big(\sum_h \|[K_h, \xi]\|_2^2\Big)^{1/2} \le c, \qquad \Big(\sum_h \|FK_h\xi\|_2^2\Big)^{1/2} \le \|F\xi\|_2 + \|F\|_{\mathrm{op}}c.$$

Let B be Bob's record of the sequence on ρ.

(a) *If a unitary V_t fixes the earlier labels and permutes the q_t labels of step t cyclically, and a_t is the sum of $c(V_t)$, of the aggregate covariance errors $(\sum_a \|(P_{s,a}V_t - V_t P_{s,\pi_s(a)})\xi\|_2^2)^{1/2}$ over $s \le t$, and of $C_Q Mc$, then $H(B_t \mid B_1, \dots, B_{t-1}) \ge \log q_t - C_Q \min(1, \sqrt{a_t})$.*

(b) *On the canonical purification, let Alice measure the transposed partitions in the same order, with record A. Then $H(B \mid A) \le C_Q \sum_t \min(1, c_t + \sum_{s<t} d_{t,s} + Mc)$, and $S(\rho) \ge H(B) - H(B \mid A)$.*

Proof. Completeness follows by summing the innermost factor. Expanding $[AB, \xi] = A[B, \xi] + [A, \xi]B$ and summing each term by one of the two completeness relations gives the centrality bound by induction; then $FK_h\xi = F\xi K_h + F[K_h, \xi]$, and $\sum_h \|F\xi K_h\|_2^2 = \|F\xi\|_2^2$ by unitality.

(a) Telescope $V_t K_h - K_{\pi h} V_t$ through the prefix; each term is a later product, a covariance defect $P_{s,a}V_t - V_t P_{s,\pi_s(a)}$, and an earlier history applied to ξ. Summing the later outcomes by completeness and the earlier histories by the second estimate bounds the aggregate by a_t; moving V_t through ξ costs $c(V_t)$. Since right multiplication by V_t preserves the norm, the prefix law $\|K_h\xi\|_2^2$ and its cyclic shift in the current label are within total variation Ca_t. Averaging over the q_t powers of the shift gives a law uniform in the current label given the past, and continuity of conditional entropy over an alphabet of fixed size gives the claim.

(b) The joint components before step t are $\sigma_{h,g} = K_h \xi K_g^*$, with h Bob's and g Alice's history. For every matrix σ, $\sum_a \|P_{t,a}\sigma - \sigma P_{t,a}\|_2^2 = 2\sum_{a \ne b} \|P_{t,a}\sigma P_{t,b}\|_2^2$, so $\sqrt{2\Pr[A_t \ne B_t]}$ is the aggregate norm of $[P_{t,a}, \sigma_{h,g}] = [P_{t,a}, K_h]\xi K_g^* + K_h[P_{t,a}, \xi]K_g^* + K_h\xi[P_{t,a}, K_g^*]$. The middle term aggregates to exactly c_t, and the two outer terms to the aggregate of $[P_{t,a}, K_h]\xi$, which telescopes to at most $\sum_{s<t} d_{t,s} + C_Q Mc$. Fano's inequality with $h(x) \le C\sqrt{x}$ and subadditivity bound $H(B \mid A)$. Finally Alice's measurement prepares on Bob an ensemble averaging to ρ, so Holevo's bound gives $I(A : B) \le S(\rho)$. □

Proof of Theorem 10.9. Put $n_i = \lfloor \log_{m_i}(1/e) \rfloor$, $N = \max_i n_i$ and $M = \sum_i n_i$, so $M, N = O(\log(1/e))$ and $em_i^{n_i} \le 1$. For each copy measure the q_i arcs of $Z_{i,j}$ for $j = n_i - 1, \dots, 0$, in this reverse order, the copies in a fixed order, with independent random offsets; this is an actual sequence of M projective measurements. By Lemma 10.10 and Lemma 10.11, $\mathbb{E}c \le CM\sqrt{(N+1)e}$, so every polynomial in M, N times $\mathbb{E}c$ tends to zero.

Bob's record. At stage (i, j) use $V = T_{i,j}$. It permutes the current arcs cyclically, and it nearly fixes the arcs of every earlier stage: earlier stages of the same copy are deeper, $l > j$, and those of other copies are controlled by the cross-copy estimates. By Lemma 10.11 and a geometric sum, $\mathbb{E}a_{i,j} \le C(em_i^{n_i - j})^{1/2} + C(N+1)^C e^{1/2}$, so by Lemma 10.12(a) and Jensen $\mathbb{E}[\log q_i - H(B_{i,j} \mid \text{earlier})] \le C(em_i^{n_i-j})^{1/4} + C(N+1)^C e^{1/4}$. The first terms form a geometric series in j, bounded uniformly in e, and the remainders sum to $o(1)$. Hence $\mathbb{E}H(B) \ge \sum_i n_i \log q_i - C$; the conditional entropies add across copies because the earlier record contains the copies already processed.

Alice's prediction. For two stages of one copy, $\mathbb{E}d \le C(em_i^{|l-j|})^{1/4} + C((N+1)e)^{1/4}$, and for different copies $\mathbb{E}d \le C((N+1)^3 e)^{1/4}$. At stage (i, j) the earlier stages of the same copy have $l > j$, and $\sum_{l>j}(em_i^{l-j})^{1/4} \le C(em_i^{n_i - j})^{1/4}$. Summing Lemma 10.12(b) gives $\mathbb{E}H(B \mid A) \le C$.

Conclusion. For every offset $S(\rho) \ge H(B) - H(B \mid A)$; averaging over offsets, which does not change ρ, $S(\rho) \ge \sum_i n_i \log q_i - C = s \log(1/e) - C$. Finally $\log \operatorname{rank} \rho \ge S(\rho)$. For the upper bound, tensor the Fourier-lamp models of the copies at precision e, with the flat state: cross-copy commutators vanish exactly, the marked involution is $-I$ in every factor, internal defects are unchanged, dimensions multiply and entropies add. □

The estimates of Lemmas 10.11 and 10.12 are written for this purpose, but each is elementary: the first is Connes' coupling with a random cut, the second uses that a sequence of projective pinchings is both trace preserving and unital, so that weighted errors can be

summed over all histories without a factor of their number. The reverse order of depths is what makes the same-copy errors geometric.

10.8 Shared dilations force uniformity

Two dilations acting on the same Y can force more than either alone. Take commuting involutions z_0, z_1, w with the row $z_0 z_1 w = 1$, a unitary y exchanging z_0 and z_1 by conjugation, and unitaries x_4, x_6 with $x_j w x_j^{-1} = w$ and $x_j y x_j^{-1} = y^j$; the relation $[x_4, x_6] = 1$ may be added. Call this source $K_{4,6}$; the marked involution is w.

Theorem 10.13 (Shared dilations 4 and 6)**.** *Every model of $K_{4,6}$ at precision e, with or without $[x_4, x_6] = 1$, satisfies $d \geq \operatorname{rank}\rho \geq c/e$ and $S(\rho) \geq \log(1/e) - C$, and there are flat models of dimension $O(1/e)$ satisfying $[x_4, x_6] = 1$ exactly. For the ρ-spectral law μ of Y, with $\delta_j = W_1(j_*\mu, \mu)$ and $\delta_S = W_1((+\frac{1}{2})_*\mu, \mu)$ in turns,*

$$W_1(\mu, h) \;\leq\; \tfrac{1}{3}(4\delta_6 + 7\delta_4 + 5\delta_S) \;\leq\; \tfrac{85}{4}e.$$

Proof. Spectral defects. The lamp swap and the row give $\|(Z_1 + Z_0)\xi\|_2 \leq 5e$, hence $\|(Z_0 Y Z_0^* + Y)\xi\|_2 \leq 12e$ after moving suffixes, and Connes' coupling with the state-change cost gives $\delta_S \leq \frac{17}{4}e$. The dilation relators give $\delta_j \leq \frac{j+3}{2}e$, so $\delta_4 \leq \frac{7}{2}e$ and $\delta_6 \leq \frac{9}{2}e$.

Inverse-root kernels. For $m \geq 1$ let D_m be the pushforward under $t \mapsto mt$, Q_m the average of the rotations by j/m, and R_m the operator choosing a uniform m-th root. Then $R_m D_m = Q_m$, $R_2 Q_2 D_2 = Q_4$, $W_1(R_m\alpha, R_m\beta) \leq \frac{1}{m} W_1(\alpha, \beta)$ (couple t with $t + \Delta$ by a shortest signed difference and the roots $(t + j)/m$ with $(t + \Delta + j)/m$), and, as R_m fixes Haar measure, $W_1(\mu, h) \leq \frac{m}{m-1} W_1(\mu, R_m\mu)$.

The arithmetic. Since $D_2 D_6 = D_3 D_4 = D_{12}$, $W_1(D_2\mu, D_3\mu) \leq 2\delta_6 + 3\delta_4$. The map D_3 commutes with the half-turn because 3 is odd and is 3-Lipschitz, so $W_1(D_2\mu, S_* D_2\mu) \leq 4\delta_6 + 6\delta_4 + 3\delta_S$. Then $W_1(\mu, Q_4\mu) \leq W_1(\mu, Q_2\mu) + \frac{1}{4} W_1(D_2\mu, S_* D_2\mu) \leq \delta_6 + \frac{3}{2}\delta_4 + \frac{5}{4}\delta_S$, and $W_1(\mu, R_4\mu) \leq W_1(\mu, Q_4\mu) + \frac{1}{4}\delta_4$. With $m = 4$ this proves the display.

Lower bounds. A law with at most d atoms is at distance at least $1/(4d)$ from Haar measure in turns, so $d \geq 1/(85e)$. Theorem 9.6 with $\nu = h$ and $\|[Y, \xi]\|_2 \leq e$ gives $S(\rho) \geq \log(1/e) - C$, and $\operatorname{rank}\rho \geq 2^{S(\rho)}$.

Upper models. For odd L put $d = 2^L$ and $N = d - 1$, which is coprime to 6; let $Y e_b = e^{2\pi i b/N} e_b$ for $0 \leq b \leq N$ (one duplicated root at $b = N$), $Z_0 e_b = e_{b \oplus 2^{L-1}}$, $Z_1 = -Z_0$, $W = -I$, and X_j the permutation matrix with $X_j e_b = e_{j^{-1} b \bmod N}$ on the cyclic labels $b < N$, where j^{-1} is the inverse of j modulo N, fixing the additional label N; equivalently $(X_j)_{b, jb} = 1$. Then $X_j^* e_b = e_{jb}$, so $X_j Y X_j^* e_b = e^{2\pi i j b/N} e_b$ and $X_j Y X_j^* = Y^j$; with the opposite convention $X_j e_b = e_{jb}$ one would get $Y^{j^{-1}}$ instead, for instance Y^2 at $N = 7$, $j = 4$. Also $X_j W X_j^* = W$ and $X_4 X_6 = X_6 X_4$ exactly, and flipping the highest bit changes b/N by $\pm(\frac{1}{2} + \frac{1}{2N})$, so the lamp swaps have operator defect at most π/N. With the flat state this is a model at precision π/N and entropy $\log d$. □

Each dilation alone permits the smaller exponent $\log_4 2 = \frac{1}{2}$ or $\log_6 2$; sharing Y forces uniformity. This is an exact elementary transport identity, not a quantitative solution of the $\times m, \times n$ rigidity problem: for the pair $4, 9$ the present bounds are $\Omega(e^{-1/2})$ and $O(e^{-1})$ (Section 18).

11 Games with a dense set of coefficients

The lamp sources and their sums realize every positive rational exponent at the source level, and more generally every number $\sum_i \log_{m_i} q_i$ with $q_i \geq 2$ a power of two dividing m_i. Compiled

into fixed games, they give the rational family, with minimum pure Schmidt rank $\Theta(\varepsilon^{-\alpha})$, the whole finite-Rényi spectrum and pure one-shot concentration (Sections 11.1 to 11.3). For every coefficient $\alpha = \frac{1}{2}\sum_i \log_{m_i} q_i$, rational or not, they give games in which the one-way distillable entanglement of every mixed near-winner is at least $\alpha \log(1/\varepsilon) - O(\log\log(1/\varepsilon))$, and the infimum demand has the sharp leading coefficient α (Section 11.4). These coefficients form a dense set that contains transcendental numbers (Section 11.5), and Section 11.6 bounds the arithmetic complexity of the coefficient of every fixed game.

11.1 The Boolean constraint family

Theorem 11.1 (Every positive rational exponent)**.** *For every positive rational $\alpha = a/b$ there is a fixed finite Boolean constraint-variable game G_α of quantum value one, with every incidence positively tested and no finite-dimensional perfect strategy, such that the minimum Schmidt rank of a pure strategy of loss at most ε is $\Theta_\alpha(\varepsilon^{-\alpha})$, and the minimum entanglement entropy of a pure strategy and the minimum entanglement of formation of a mixed strategy are both $\alpha \log(1/\varepsilon) + O_\alpha(1)$. The game is fixed before ε is chosen.*

Proof. Take $k = 2a$ copies of the binary lamp source $L_{2,2^b}$, so that Theorem 10.9 applies with $s = k/b = 2\alpha$, and let G_α be the game of the universal compiler (Theorem 5.6) for this presentation, with contexts of at most three variables. A strategy of loss ε gives, by Theorem 4.9(a), a model of the constraint system at precision $C\sqrt{\varepsilon}$ on Bob's Schmidt support with Bob's reduced state, and the lower map of Theorem 5.6 turns it into a model of the source with the *same* state at precision $C'\sqrt{\varepsilon}$. By Theorem 10.9 its rank is at least $c\varepsilon^{-s/2} = c\varepsilon^{-\alpha}$ and its entropy at least $\alpha \log(1/\varepsilon) - C$. Conversely, the flat tensor-product models at precision e, lifted by the upper map with the state $I_{12}/12 \otimes \rho$ and returned by Theorem 4.9(b), give pure strategies of loss $K_3 C^2 e^2$, Schmidt rank $O(e^{-s})$ and entropy $s \log(1/e) + \log 12 + O(1)$. Taking e of order $\sqrt{\varepsilon}$ proves the pure statements. For the entanglement of formation, the pure lower bound $f(\ell) = \alpha \log(1/\ell) - C$ is convex and decreasing; it holds for every pure strategy of loss $\ell > 0$ after enlarging C, and no pure strategy has $\ell = 0$, since the source has no finite-dimensional exact model with the marked involution $-I$. For a decomposition $\sigma = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ with the same measurements the losses average to at most ε, and Jensen's inequality gives $\sum_j p_j S((\psi_j)_B) \geq f(\varepsilon)$. The upper bound is attained by the pure strategies above. □

This is a statement about a Boolean constraint game and needs no linear compiler. The mixed-state E_{D} statement is Theorem 11.9; it does not follow from the bound on E_{F}.

11.2 Linear system games

Theorem 11.2 (The one-copy linear family)**.** *For every $q = 2^r$ and every m divisible by q there is a fixed finite linear system game of quantum value one whose minimum pure Schmidt rank at loss ε is $\Theta(\varepsilon^{-\alpha})$, $\alpha = \frac{1}{2}\log_m q$, and whose minimum pure entanglement entropy and minimum mixed entanglement of formation are $\alpha \log(1/\varepsilon) + O(1)$. In particular every positive rational $\alpha \leq \frac{1}{2}$ occurs.*

Proof. The lamp source $L_{q,m}$ is an extended homogeneous-linear-plus-conjugacy presentation in the sense of Slofstra and Vidick [134, §6, (a)–(d)]: its involutions are $z_0, \ldots, z_{q-1}, w$, the single set $V = \{z_0, \ldots, z_{q-1}, w\}$ carries the commuting row $z_0 \cdots z_{q-1} w = e$, the relations $y z_j y^{-1} = z_{j+1}$ and $x w x^{-1} = w$ have form (d)(ii), and $xyx^{-1} = y^m$ has form (d)(iii) with $y < x$. Its marked involution w is nontrivial in approximate representations, by the Fourier-lamp models. Proposition 5.10, with the $l = 2$ non-involutive generators $y < x$ and $N_G = 16$, therefore applies to the solution group Γ of the balanced construction: every strategy of loss

ε has Bob's nonzero spectrum in $\Lambda_{L_{q,m}}(C_1\sqrt{\varepsilon})$, so by Theorem 10.2 its rank is at least $c\varepsilon^{-\alpha}$ and its entropy at least $\alpha\log(1/\varepsilon) - C$; and flat models of $L_{q,m}$ of dimension $O(e^{-2\alpha})$ give strategies of Schmidt rank $16 \cdot O(e^{-2\alpha})$ and loss $O(e^2)$. The formation statement follows as in Theorem 11.1. For $\alpha = a/b \le \frac{1}{2}$ take $r = 2a$ and $m = 2^b$. □

We also use a restricted linear compiler of our own, which eliminates the non-involutive generators in order (a generator Y becomes the product fv of two new involutions, and a power relation $UYU^{-1} = Y^m$ the conjugacy $UvU^{-1} = p_m$ with the involutive palindrome $p_m = v(fv)^{m-1}$), adds a parity cover with first-negative corner selectors, which is the multi-corner map of Proposition 5.5, and finishes with Slofstra's gadgets. For one copy it gives amplification 16 and was used for the explicit exports below. The balanced construction behind Proposition 5.10 has the same architecture, with Slofstra's conversions, and the same amplification 2^{l+2} for its own games; it does not certify the specifications below.

Theorem 11.3 (The restricted compiler for several copies; conditional)**.** *The restricted compiler applies to the input grammar of homogeneous commuting-involution rows, involutive conjugacies, noninvolutive conjugation of an involution, and strictly later-on-earlier positive-power actions (including power one), with a shared central negative involution; it does not compile arbitrary signed rows or arbitrary presentations. For the k-copy sources of Theorem 10.9, with $2k$ ordered eliminations, the recorded maximum arity is $\max\{3, \max_i(q_i + 1), 2^{2k-1}+1\}$, the amplification 2^{2k+2}, and the entropy overhead $2k+2$; the precision constants are computable from the finite maps, and no uniform constant for the whole family is asserted.* Conditional on a self-contained proof of the family compiler for general k, *the compiler itself converts these sources into fixed linear system games with these specifications and the profiles of Theorem 11.1 for every positive rational α.*

Linear system games with these profiles exist unconditionally, by the route of Corollary 11.6 below, through Proposition 5.10; that route does not certify the arity, the amplification, the algorithm or the exports of the restricted compiler in Theorem 11.3, which keep their condition.

The return step must use arity-dependent constants: for a context of k variables, Theorem 4.5 with $s = k$, or the boolean return with $K_k D_k^2$, where $D_k = \max\{1, [1 + k(k-1)(3k+1)/4]/2\}$ converts parity errors into forbidden-product errors (Remark 4.11). At arity nine $K_9 = 138496$ and $K_9 D_9^2 = 8829985600$; the ternary constant 111 must never be used there.

Remark 11.4 (Exported objects; checked). Three finite objects were exported and checked exactly. K4 has 5206 game variables and 4325 parity rows, of arity three, and L4M8 ($q = 4$, $m = 8$) has 6798 game variables and 5653 rows, of maximum arity five; both have amplification 16, their exponents are $\frac{1}{4}$ and $\frac{1}{3}$, and their solution-group presentations have one more generator, the central J. A two-copy object has 199129 game variables, again without J, 164608 rows, maximum arity nine, amplification 64 and exponent one. The script `verify/check_linear_games.py` rebuilds K4 and L4M8 from their sources and checks exactly, in the negative source algebra, every square, parity row and row commutator and the restriction of each conjugated source word to its generator on its block. The exported L4M8 and two-copy games are archived in `verify/artifacts`: the archived L4M8 file is checked exactly against its source presentation, every square, row and row commutator, with the normal-form engine of `check_linear_games.py`, and the two-copy file by its own exact checker; complete Fourier-lamp strategies were played on every context of L4M8 at four sizes, with observed rank exponents 0.318, 0.331, 0.333, and on every context and raw outcome of the two-copy game at two sizes. This is evidence for these finite objects, not a derivation of every k.

Remark 11.5 (Several copies are an extended presentation). The source $L_{\mathbf{q},\mathbf{m}}$ of Theorem 10.9 is an extended homogeneous-linear-plus-conjugacy presentation [134, §6, (a)–(d)], after a Tietze transformation, with the noninvolutive generators ordered $y_1, \dots, y_k, x_1, \dots, x_k$. Its involutions are the lamps of all copies and the shared w; the rows $z_0^{(i)} \cdots z_{q_i-1}^{(i)} w = e$, one set of the collection in (c) for each copy, carry the within-copy commutations; and within a copy the remaining relations have forms (d)(ii) and (d)(iii) as in the proof of Theorem 11.2, $x_i y_i x_i^{-1} = y_i^{m_i}$ being a later-on-earlier positive power. The cross-copy commutators are of three kinds. Two involutions u, v commute exactly when $uvu = v$, an allowed involutive conjugacy (d)(i); a two-element row $uv = e$ of form (c) would instead impose $u = v$. A noninvolutive generator g commutes with an involution u through the allowed action $gug^{-1} = u$ of form (d)(ii). Two noninvolutive generators commute through the later-on-earlier power-one relation $g_i g_j g_i^{-1} = g_j^1$, $j < i$, of form (d)(iii). The shared central w stays one involution, and tensor products of the negative Fourier-lamp models respect that identification. Proposition 5.10, with $l = 2k$ non-involutive generators and its multi-corner lower map, and Theorem 10.9 therefore give an unconditional linear realization of every positive rational exponent, with flat amplification $N_G = 2^{2k+2}$ and fixed computable presentation-dependent precision constants. Nothing here uses that m_i is a power of two, which is how Theorem 11.9 obtains linear games for every coefficient of Corollary 11.11.

Corollary 11.6 (Linear games with every positive rational exponent)**.** *For every positive rational $\alpha = a/b$ there is a fixed finite linear system game $\widetilde{G}_\alpha$ of quantum value one, with every incidence positively tested and no finite-dimensional perfect strategy, such that the minimum Schmidt rank of a pure strategy of loss at most ε is $\Theta_\alpha(\varepsilon^{-\alpha})$, the minimum entanglement entropy of a pure strategy and the minimum entanglement of formation of a mixed strategy are $\alpha \log(1/\varepsilon) + O_\alpha(1)$, the pure-state statements of Theorem 11.7 hold for $\widetilde{G}_\alpha$, and there is $\varepsilon_\alpha > 0$ such that, separately in both directions, for $0 < \varepsilon < \varepsilon_\alpha$,*

$$d^{\rightarrow}_{\widetilde{G}_\alpha}(\varepsilon) = \alpha \log(1/\varepsilon) + O_\alpha\big(\log\log(1/\varepsilon)\big), \qquad d^{\leftarrow}_{\widetilde{G}_\alpha}(\varepsilon) = \alpha \log(1/\varepsilon) + O_\alpha\big(\log\log(1/\varepsilon)\big).$$

This is not an equality of the two functions at each loss. The corollary concerns existence only: it does not certify the specifications of Theorem 11.3.

Proof. Take $k = 2a$ binary copies, $q_i = 2$ and $m_i = 2^b$, so that $s = k/b = 2\alpha$ in Theorem 10.9, and let $L = L_{\mathbf{q},\mathbf{m}}$ with marked involution w. By Remark 11.5, L is an extended homogeneous-linear-plus-conjugacy presentation after a Tietze transformation, which changes precisions by constant factors and keeps the state (Proposition 5.3). The involution w is central, and it is nontrivial in approximate representations, since tensor products of the Fourier-lamp models of Example 10.7 have $W = -I$ at every precision. So the balanced construction before Proposition 5.10, with $l = 2k = 4a$ non-involutive generators, gives a binary linear system (A, b), and Proposition 5.10 applies to $\widetilde{G}_\alpha = G(A, b)$ with either question distribution of Definition 4.1 and $N_G = 2^{4a+2}$; both test every incidence.

Pure strategies. By Proposition 5.10(a) Bob's nonzero spectrum lies in $\Lambda_L(C_1\sqrt{\varepsilon})$, so Theorem 10.9 gives Schmidt rank at least $c\varepsilon^{-\alpha}$ and entropy at least $\alpha \log(1/\varepsilon) - C$. The flat tensor models of Theorem 10.9, of dimension $O(e^{-s})$ with $W = -I$, give $\mathrm{hlp}_L(w; 2, e) = O(e^{-s})$, and Proposition 5.10(b) turns them into strategies on maximally entangled states of Schmidt rank $O(\varepsilon^{-\alpha})$ and entropy $\alpha \log(1/\varepsilon) + O(1)$. No finite-dimensional strategy is perfect, since Theorem 10.9 excludes exact finite models with $W = -I$; the formation statement follows as in Theorem 11.1. The proof of Theorem 11.7 uses only the energy functional of the source on Bob's original state, which the transfer of Proposition 5.10(a) supplies at precision $O(\sqrt{\varepsilon})$, these upper strategies, and a product strategy of positive loss; it applies verbatim.

Mixed strategies. The lower map of Proposition 5.10 is the corner map of Proposition 5.5 with 2^{4a-1} corners, a unital $*$-homomorphism given by fixed polynomials in the variables, the

letter J evaluating to -1. So Proposition 6.1, in the form noted after it for linear system games and corner maps, supplies the actual source interface on the original mixed state, and the usable-entanglement law in both directions is the case $k = 2a$, $q_i = 2$, $m_i = 2^b$ of Theorem 11.9 below, for the game $\widetilde{G}_\alpha$. □

11.3 The Rényi spectrum and pure concentration

For a density matrix ρ and $1 < \beta < \infty$ write $H_\beta(\rho) = \frac{1}{1-\beta} \log \operatorname{Tr} \rho^\beta$, $H_\infty(\rho) = -\log\|\rho\|_{\mathrm{op}}$, and $H^\delta_{\min}(\rho) = \sup\{-\log\|\sigma\|_{\mathrm{op}} : P(\rho,\sigma) \le \delta\}$ with the purified distance $P = \sqrt{1-F^2}$ between normalized states.

Theorem 11.7 (Rényi spectrum)**.** *Fix a rational $\alpha > 0$ and the game G_α of Theorem 11.1. Over pure strategies of loss at most ε, with $\rho = \rho_B$,*

$$\inf H_\beta(\rho) = \alpha \log(1/\varepsilon) + O(1) \qquad (0 \le \beta \le 1),$$
$$\inf H_\beta(\rho) = \min\Big\{\alpha, \frac{\beta}{\beta-1}\Big\} \log(1/\varepsilon) + O_\beta(1) \qquad (1 < \beta < \infty),$$

the finite critical order included; the constants are not uniform as $\beta \to \infty$. For H_∞ the coefficient is $\min(\alpha, 1)$ with $O(1)$ remainder when $\alpha \ne 1$; at $\alpha = 1$,

$$\log(1/\varepsilon) - \log(1 + \log(1/\varepsilon)) - O(1) \;\le\; \inf H_\infty(\rho) \;\le\; \log(1/\varepsilon) + O(1).$$

For fixed $0 < \delta < 1$, $\inf H^\delta_{\min}(\rho) = \alpha \log(1/\varepsilon) + O_\delta(1)$. For every fixed $0 < \eta < 1$, every pure strategy of loss at most ε admits a state-dependent heralded LOCC protocol that succeeds with probability at least $1 - \eta$ and then outputs an exact Φ_M with $\log M \ge \alpha \log(1/\varepsilon) - O_\eta(1)$; this uses the Schmidt basis of the state and is not the protocol of Theorem 7.7. The same holds for the linear games of Theorem 11.2 and of Corollary 11.6.

The proof turns the rank profile into a functional inequality. Fix the source unitaries of an arbitrary assignment and define, for positive A,

$$E(A)^2 = \sum_R \|(R - I)A\|_2^2 + \sum_V \|[V, A]\|_2^2 + \|(W + I)A\|_2^2,$$

summing over the relators R, the generators V and the marked involution W. A model on $\xi = \rho^{1/2}$ at precision e has $E(\xi) \le Ce$, and the game transfer keeps ρ and gives $E(\xi) \le C\sqrt{\varepsilon}$.

Lemma 11.8 (Faber–Krahn and layers)**.** *Let $s = 2\alpha$. There is $c > 0$ with $E(A)^2 \ge c \operatorname{rank}(A)^{-2/s} \|A\|_2^2$ for every positive A and every assignment. With $f_l(t) = \min((t - 2^l)_+, 2^l)$ for $l \in \mathbb{Z}$, $\sum_l E(f_l(A))^2 \le E(A)^2$, and $E((A - tI)_+) \le E(A)$ for $t \ge 0$.*

Proof. Apply Theorem 10.9 to the density $A^2/\|A\|_2^2$ with the same unitaries, at precision $E(A)/\|A\|_2$; lowering c covers large precisions, and $E(A) = 0$ is impossible for $A \ne 0$ since the rank bound would then hold at every precision. For the layers, $\sum_l f_l(t)^2 \le t^2$ and $\sum_l |f_l(t) - f_l(u)|^2 \le (t-u)^2$, the $f_l(t)$ being the lengths of the disjoint intervals $[2^l, 2^{l+1}] \cap [0, t]$. So $\sum_l \|L f_l(A)\|_2^2 = \operatorname{Tr}(L^* L \sum_l f_l(A)^2) \le \|LA\|_2^2$, and in an eigenbasis of A, $\sum_l \|[V, f_l(A)]\|_2^2 = \sum_{i,j} |V_{ij}|^2 \sum_l |f_l(a_j) - f_l(a_i)|^2 \le \|[V, A]\|_2^2$. The cutoff $(t-u)_+$ is 1-Lipschitz and vanishes at zero, which gives the last claim in the same way. □

Proof of Theorem 11.7. Schatten bounds. First let $s > 2$, $\vartheta = 2/s$ and $p = 2s/(s-2)$. With $r_l = \operatorname{rank} f_l(A)$ and $b_l = 2^{pl} r_l$, the lemma gives $E(f_l(A))^2 \ge c 2^{2l} r_{l+1} r_l^{-\vartheta} = c 2^{-p} b_{l+1} b_l^{-\vartheta}$, since $2 - p + p\vartheta = 0$; with $T = \sum_l b_l$ and $b_l \le T$, $\sum_l b_{l+1} b_l^{-\vartheta} \ge T^{1-\vartheta}$, and dyadic layer cake gives $T \asymp \operatorname{Tr} A^p$. Hence $\|A\|_p \le C\, E(A)$, the strong endpoint, without a logarithm. For any

$s > 0$ and $p > 2$ with $t = 2p/(p-2) > s$, add the mass term $\lambda\|B\|_2^2$ to the energy; since $c\,\mathrm{rank}^{-2/s} + \lambda \geq c^a \lambda^{1-a}\,\mathrm{rank}^{-2/t}$ with $a = s/t$, the endpoint with parameter t and the choice $\lambda = E(A)^2/\|A\|_2^2$ give $\|A\|_p \leq C\,E(A)^a \|A\|_2^{1-a}$. Above the endpoint use monotonicity of Schatten norms. With $A = \xi$, $\|A\|_2 = 1$ and $E(A) \leq e$, this gives $\mathrm{Tr}\,\rho^\beta \leq C e^{\min(s(\beta-1),2\beta)}$, that is $H_\beta(\rho) \geq \min(s, 2\beta/(\beta-1))\log(1/e) - C$, and with $e \asymp \sqrt{\varepsilon}$ the lower bounds for $1 < \beta < \infty$. For $0 \leq \beta \leq 1$, $H_\beta \geq H_1$, and Theorem 11.1 gives the coefficient α.

Min-entropy. If $s > 2$, $\|A\|_{\mathrm{op}} \leq \|A\|_p \leq Ce$, so $\|\rho\|_{\mathrm{op}} \leq Ce^2$. If $s < 2$, normalize $\|A\|_2 = 1$ and let $a_1 \geq a_2 \geq \cdots$ be the eigenvalues; $B = (A - a_{2r})_+$ has rank below $2r$, $\|B\|_2 \geq \sqrt{r}(a_r - a_{2r})$ and $E(B) \leq e$, so $a_r - a_{2r} \leq Ce r^{1/s-1/2}$; telescoping over dyadic r up to $R \asymp e^{-s}$ and $a_R \leq R^{-1/2}$ give $a_1 \leq Ce^{s/2}$, that is $\|\rho\|_{\mathrm{op}} \leq Ce^s$. If $s = 2$, normalize $\|A\|_2 = 1$, let $M = \|A\|_{\mathrm{op}}$, and suppose $M \geq E(A)$; otherwise $M^2 < E(A)^2$ and there is nothing to prove. Put $K_0 = \lceil \log(4/M^2) \rceil \geq 2$, $\Delta = M/(4K_0)$, $t_j = M/2 + j\Delta$, $h_j = \mathrm{rank}(A > t_j)$ and $F_j = \min\{(A - t_j)_+, \Delta\}$ for $0 \leq j < K_0$. Then $h_0 \leq \|A\|_2^2/(M/2)^2 = 4/M^2$ and $h_{K_0} \geq 1$, since $t_{K_0} = 3M/4 < M$. The bands $[t_j, t_j + \Delta]$ are disjoint, so, as for the layers of Lemma 11.8, $\sum_j E(F_j)^2 \leq E(A)^2$; and $F_j = \Delta$ on the eigenvalues above t_{j+1}, so $\|F_j\|_2^2 \geq \Delta^2 h_{j+1}$, while $\mathrm{rank}\,F_j = h_j$. The $s = 2$ inequality $E(F)^2 \geq c\,\mathrm{rank}(F)^{-1}\|F\|_2^2$ and the arithmetic-geometric mean inequality give

$$E(A)^2 \geq c\Delta^2 \sum_{j<K_0} \frac{h_{j+1}}{h_j} \geq c\Delta^2 K_0 \Big(\frac{h_{K_0}}{h_0}\Big)^{1/K_0} \geq \frac{cM^2}{16K_0}\Big(\frac{M^2}{4}\Big)^{1/K_0} \geq \frac{cM^2}{32K_0},$$

and $K_0 = O(1 + \log(1/E(A)))$ because $M \geq E(A)$. So $\|\rho\|_{\mathrm{op}} = M^2 \leq CE(A)^2[1 + \log(1/E(A))]$, that is $\|\rho\|_{\mathrm{op}} \leq Ce^2(1 + \log(1/e))$. With $e \asymp \sqrt{\varepsilon}$ these are the lower bounds for H_∞; at $s = 2$, substituting $e = C_G\sqrt{\varepsilon}$ gives $H_\infty(\rho) \geq \log(1/\varepsilon) - \log(1 + \log(1/\varepsilon)) - O(1)$, a single logarithmic-logarithmic loss. The $O(1)$ endpoint remainder is open, and the constants are not uniform in the Rényi order.

Upper bounds. Take a pure strategy of loss at most $\varepsilon/2$ and Schmidt rank $D \leq C\varepsilon^{-\alpha}$, and a deterministic product strategy of loss $\ell_{\mathrm{bad}} > 0$; such a positive loss exists because there is no perfect finite strategy. With $\delta = \varepsilon/(2\ell_{\mathrm{bad}})$ and orthogonal local flags on both sides, $\sqrt{1-\delta}|0,0\rangle|\psi_{\mathrm{good}}\rangle + \sqrt{\delta}|1,1\rangle|a\rangle|b\rangle$ is a pure strategy of loss at most ε whose squared Schmidt coefficients are $(1-\delta)\lambda_j$ and δ. So $\mathrm{Tr}\,\rho^\beta \geq (1-\delta)^\beta D^{1-\beta} + \delta^\beta \geq c(\varepsilon^{\alpha(\beta-1)} + \varepsilon^\beta)$ and $\|\rho\|_{\mathrm{op}} \geq \max((1-\delta)/D, \delta)$, which give the upper bounds.

Smoothing and concentration. Choose $\beta > 1$ with $\alpha \leq \beta/(\beta-1)$, so $\mathrm{Tr}\,\rho^\beta \leq C\varepsilon^{\alpha(\beta-1)}$. For $\eta \in (0,1)$ let $t = (C/\eta)^{1/(\beta-1)}\varepsilon^\alpha$ and Q the spectral projection of ρ onto eigenvalues at most t. Then $\mu = \mathrm{Tr}(I-Q)\rho \leq t^{1-\beta}\,\mathrm{Tr}\,\rho^\beta \leq \eta$, and $\sigma = Q\rho Q/(1-\mu)$ has $F(\rho,\sigma) = \sqrt{1-\mu}$ and $\|\sigma\|_{\mathrm{op}} \leq t/(1-\eta)$; with $\eta = \delta^2$ this is the lower bound for $H_{\min}^\delta$. For the upper bound, a strategy of rank D and any σ with $P(\rho,\sigma) \leq \delta$ have $\mathrm{Tr}(\Pi\sigma) \geq 1 - \delta^2$ on the support Π, so $\|\sigma\|_{\mathrm{op}} \geq (1-\delta^2)/D$. For concentration, project locally onto Q in the Schmidt basis, which succeeds with probability $1 - \mu \geq 1 - \eta$, and convert the resulting pure state, whose largest Schmidt weight is at most $1/M$ with $M = \lfloor (1-\eta)/t \rfloor$, to Φ_M deterministically by Nielsen's theorem [109]. Schmidt rank cannot increase in a pure branch, so $M \leq D$ for the upper scale. □

The weight in the largest r Schmidt coefficients is controlled as well: by Hölder, $L_r \leq C(r\varepsilon^\alpha)^{(\beta-1)/\beta}$, so for every $r = o(\varepsilon^{-\alpha})$ it tends to zero uniformly over near-winning pure strategies. The functional-inequality framework is that of Bakry, Coulhon, Ledoux and Saloff-Coste and of Rumin [7, 126]; one-shot concentration is due to Vidal, Lo–Popescu and Buscemi–Datta [144, 90, 25]. A large rank alone is not a large entanglement: it could sit in a negligible tail, and the theorem excludes that at the certified scale. For the source of Theorem 10.13, $s = 1$ and the same argument gives $\|\rho\|_{\mathrm{op}} \leq Ce$, so every Rényi entropy, H_∞ included, has coefficient one there.

11.4 Usable entanglement with the sharp coefficient

The coefficients of the lamp sums need not be rational. For $k \geq 1$ and $i = 1, \ldots, k$ fix integers $r_i \geq 1$ and $m_i \geq 2$ with $2^{r_i} \mid m_i$, put $q_i = 2^{r_i}$, let $L_{\mathbf{q},\mathbf{m}}$ be the source of Theorem 10.9, and put

$$\alpha = \frac{1}{2}\sum_{i=1}^{k} \log_{m_i} q_i = \frac{1}{2}\sum_{i=1}^{k} \frac{r_i}{\log m_i}.$$

Let $G_{\mathbf{q},\mathbf{m}}$ be the Boolean constraint game of Theorem 5.6 for $L_{\mathbf{q},\mathbf{m}}$, and $\widetilde{G}_{\mathbf{q},\mathbf{m}}$ the linear system game that Remark 11.5 and Proposition 5.10 give for it, with either question distribution of Definition 4.1; the remark uses only the form of the relators, not the values of q_i and m_i. For $\alpha = a/b$, the choice of $k = 2a$ copies of $L_{2,2^b}$ gives the games G_α and $\widetilde{G}_\alpha$ of Theorem 11.1 and Corollary 11.6, and one copy gives the games of Theorem 11.2.

Theorem 11.9 (Sharp coefficients on the original state)**.** *Let G be $G_{\mathbf{q},\mathbf{m}}$ or $\widetilde{G}_{\mathbf{q},\mathbf{m}}$. Then G has quantum value one, every incidence positively tested and no finite-dimensional perfect strategy, and there are constants $\varepsilon_G, C_G > 0$ such that, for $0 < \varepsilon < \varepsilon_G$:*

(a) the minimum Schmidt rank of a pure strategy of loss at most ε is $\Theta_G(\varepsilon^{-\alpha})$, and the minimum entanglement entropy of a pure strategy and the minimum entanglement of formation of a mixed strategy are $\alpha \log(1/\varepsilon) + O_G(1)$;

(b) every finite-dimensional mixed strategy of loss at most ε satisfies, separately in both directions,

$$\min\{E_{\mathrm{D}}^{A\to B}(\rho_{AB}), E_{\mathrm{D}}^{B\to A}(\rho_{AB})\} \;\geq\; \alpha \log(1/\varepsilon) - O_G\big(\log\log(1/\varepsilon)\big),$$

and there are pure maximally entangled strategies of loss at most ε whose two directional rates are $\alpha \log(1/\varepsilon) + O_G(1)$; so $d_G^{\rightarrow}(\varepsilon)$ and $d_G^{\leftarrow}(\varepsilon)$ are both $\alpha \log(1/\varepsilon) + O_G(\log\log(1/\varepsilon))$, and an $O(1)$ lower remainder is not proved;

(c) explicitly, with $e = C_G\sqrt{\varepsilon}$, $x = \log(1/e)$ and

$$L_i = \Big\lfloor \frac{x - 32\log(2+x)}{\log m_i} \Big\rfloor, \qquad D = \prod_i q_i^{L_i},$$

one has $L_i \geq 1$, and every finite-dimensional mixed strategy of loss at most ε admits strategy-dependent product local isometries extracting Φ_D with vector error at most $C'_G(2+x)^{-3}$, the environment included in the junk.

The restriction $\varepsilon < \varepsilon_G$ ensures that the depths L_i are positive, that the error bounds of the proof are in range, and that $\log\log(1/\varepsilon)$ is defined. For rational α the theorem contains the usable-entanglement laws of the games G_α and $\widetilde{G}_\alpha$ of Theorem 11.1 and Corollary 11.6, and it adds that law to the one-copy games of Theorem 11.2.

Proof. The game and the pure and formation laws. These are proved as in Theorem 11.1 and Corollary 11.6, with Theorem 10.9 for $s = 2\alpha$ in place of the rational choice. For $G_{\mathbf{q},\mathbf{m}}$, Theorem 4.9(a) and the lower map of Theorem 5.6 give a model of $L_{\mathbf{q},\mathbf{m}}$ with Bob's reduced state at precision $C\sqrt{\varepsilon}$; for $\widetilde{G}_{\mathbf{q},\mathbf{m}}$, Proposition 5.10(a) puts Bob's nonzero spectrum in $\Lambda_{L_{\mathbf{q},\mathbf{m}}}(C\sqrt{\varepsilon})$. In both cases Theorem 10.9 gives Schmidt rank at least $c\varepsilon^{-\alpha}$ and entropy at least $\alpha \log(1/\varepsilon) - C$. The flat tensor-product Fourier-lamp models at precision $c\sqrt{\varepsilon}$, lifted by the upper map of Theorem 5.6 with amplification 12 and returned by Theorem 4.9(b), or turned into strategies by Proposition 5.10(b) with amplification $N_G = 2^{2k+2}$, give strategies of loss at most ε on maximally entangled states of Schmidt rank $O(\varepsilon^{-\alpha})$ and entropy $\alpha \log(1/\varepsilon) + O(1)$, whose two directional rates equal their entropy [12]. They show that G has quantum value one,

and no finite-dimensional strategy is perfect, since Theorem 10.9 excludes exact finite models with $W = -I$. The formation statement follows by Jensen's inequality as in Theorem 11.1. Both games test every incidence.

The source on the actual state. Fix a strategy of loss at most ε, dilate it locally, and let Ψ purify the dilated state. Proposition 6.1, applied to the lower map of Theorem 5.6, or, for $\widetilde{G}_{\mathbf{q},\mathbf{m}}$, to the corner map in the form noted after it, gives for every copy i Bob unitaries X_i, Y_i and Alice unitaries X_i^A, Y_i^A such that $X_i Y_i X_i^* = Y_i^{m_i}$ on Bob and $X_i^{A*} Y_i^A X_i^A = (Y_i^A)^{m_i}$ on Alice hold with residual at most $e = C_G\sqrt{\varepsilon}$, all cross-copy generator commutators have residual at most e, and all mirrors, inverse mirrors and square-root-state commutators are at most e; this holds on the whole original state. For the rotations apply the paragraph on selector polynomials after Proposition 6.1 to the selector T_i of Lemma 10.3 for copy i: in the negative quotient $T_i^* T_i = T_i T_i^* = 1$ and $T_i y_i = \zeta_i y_i T_i$, $\zeta_i = e^{2\pi i/q_i}$, and T_i commutes with every generator of another copy, being a polynomial in the lamps of copy i. This gives Bob unitaries U_i and Alice unitaries U_i^A, with mirrors, inverse mirrors and square-root-state commutators $O(e)$,

$$\|(U_i Y_i U_i^* - \zeta_i Y_i)\Psi\| = O(e), \qquad \|(U_i^A Y_i^A U_i^{A*} - \zeta_i^{-1} Y_i^A)\Psi\| = O(e),$$

the inverse phase on Alice because she realizes the opposite algebra, and commutators $O(e)$ of U_i and of U_i^A with the generators of the other copies on the same party. No sector is selected, and the U_i need not have finite order. We enlarge C_G to absorb the constants. Choose ε_G so small that $L_i \geq 1$ and that the error bounds η, δ and γ below lie in the ranges in which they are used, in particular $\eta \leq 1$ and $\gamma^2 < \frac{1}{2}$ in (2.4); this is possible because each of them tends to zero with ε.

Powers without an exponential square. For opposite-party unitaries, $\|(R_A^n - S_B^n)\Psi\| \leq n\|(R_A - S_B)\Psi\|$ (Lemma 2.11(c)); this is not asserted for two operators of one party. On each party and for each copy, with $S = X_i$ on Bob or $S = X_i^{A*}$ on Alice and Y the Y_i of that party, put $F_d = S^d Y S^{-d} - Y^{m^d}$, $m = m_i$. Then

$$F_{d+1} = S F_d S^{-1} + ((SYS^{-1})^{m^d} - (Y^m)^{m^d}).$$

Moving the one-letter suffix S^{-1} of the first term to its actual mirror costs $2e$. For the second term compare SYS^{-1} with the opposite-party mirror of Y^m, at cost $(m+1)e$, raise this opposite-party comparison to the power m^d, and compare the mirror's power m^{d+1} back with $Y^{m^{d+1}}$ on the same party, at cost $m^{d+1}e$. So $\|F_{d+1}\Psi\| \leq \|F_d\Psi\| + 2e + (2m+1)m^d e$, and $\|F_d\Psi\| \leq C_m m^d e$; no term $m^{2d}e$ appears. In the same way, raising the rotation relation across the parties, $\|(U_i Y_i^n U_i^* - \zeta_i^n Y_i^n)\Psi\| \leq C_G n e$ for $n \geq 1$, with ζ_i^{-n} on Alice.

Common roots and controls. For copy i put $N_i = L_i - 1$ and, on Bob,

$$Y_{i,j} = X_i^{-j} Y_i X_i^j, \qquad V_{i,j} = X_i^{-j} U_i X_i^j, \qquad R_{i,j} = Y_{i,N_i}^{m_i^{N_i - j}} \qquad (0 \leq j \leq N_i),$$

with the reversed words on Alice, for example $Y_{i,j}^A = (X_i^A)^j Y_i^A (X_i^A)^{-j}$; put $N = \max_i N_i$ and $M = \sum_i L_i$. All $R_{i,j}$ of one party and copy are powers of one unitary and commute exactly. Conjugating F_d by a word of length $O(N+1)$, $\|(Y_{i,j} - Y_{i,l}^{m_i^{l-j}})\Psi\| \leq C_G(m_i^{l-j} + N + 1)e$ for $j \leq l$, and in particular

$$\|(R_{i,j} - Y_{i,j})\Psi\| \leq C_G(m_i^{N_i - j} + N + 1)e$$

on each party, likewise for adjoints. With the mirror of $Y_{i,j}$, which costs $O((N+1)e)$ by Lemma 2.11(a), this gives $\|(R_{i,j}^A - R_{i,j}^B)\Psi\| \leq C_G(m_i^{N_i - j} + N + 1)e$, and Lemma 2.11(d) gives the square-root-state commutators. For $j \leq l$, compare $R_{i,j}$ with $Y_{i,l}^n$, $n = m_i^{l-j}$, and

conjugate the rotation estimate for Y_i^n by the native word X_i^{-l}, moving its suffix through actual mirrors:

$$\|(V_{i,l}R_{i,j}V_{i,l}^* - \zeta_i^n R_{i,j})\Psi\| \le C_G(m_i^{L_i} + N + 1)e.$$

Since q_i divides m_i, $\zeta_i^n = 1$ for $j < l$, and $\zeta_i^n = \zeta_i$ for $j = l$: the control of stage (i,l) fixes the earlier roots of its copy and rotates its own by $\frac{1}{q_i}$ of a turn, by $-\frac{1}{q_i}$ on Alice. The same holds for the powers $V_{i,l}^h$, $|h| \le Q = \max_i q_i$, at cost C_Q times larger. Words of different copies have length $O(N+1)$ and commute after $O((N+1)^2)$ swaps of generators or controls, each with a suffix of length $O(N+1)$, so their commutators on Ψ are $O((N+1)^3 e)$. Replacing a native $Y_{i,j}$ by $R_{i,j}$ costs the estimate above; when the replaced factor stands to the left of an operator of the other copy, that operator is first moved to the other party through its actual mirror, so the two large powers are never multiplied. Hence every mirror, inverse mirror, square-root-state commutator, triangular rotation and cross-copy commutator among the R, the V and their powers $|h| \le Q$ has residual at most

$$\eta = C_G\Big(e \max_i m_i^{L_i} + e(N+1)^3\Big),$$

on both parties. No relation between different controls is used, and no term $e\prod_i m_i^{L_i}$ appears.

Spectral cuts. Order the stages (i,j) copy by copy and with increasing j within a copy, M in all. For each stage choose an offset $\vartheta_{i,j}$ uniform modulo $1/q_i$, independent over the stages and the same on both parties, and let $P_{i,j,a}^P$, $a \in \mathbb{Z}_{q_i}$, be the spectral projections of $R_{i,j}^P$ onto the arcs of Lemma 10.11. Since multiplication by ζ_i moves each arc to the next, the control $V_{i,j}^B$ approximately sends the label a to $a-1$ and $V_{i,j}^A$ sends a to $a+1$, which is the orientation of Lemma 6.3 on Alice. Within a copy the partitions commute exactly. For the mirrors, let ν be the joint spectral measure of the commuting unitaries R^A, R^B of two parties in Ψ; an offset uniform modulo $1/q$ separates two points at distance $d \le \frac{1}{2}$, in turns, with probability at most qd, and $d \le \frac{1}{4}|z-w|$, so, as in Lemma 17.2,

$$\mathbb{E}\sum_a \|(P_a(R^A) - P_a(R^B))\Psi\|^2 \le 2q\int d\,d\nu \le \frac{q}{2}\|(R^A - R^B)\Psi\|.$$

On Alice, with $\xi = \rho_A^{1/2}$ and vector norms on Ψ equal to Hilbert–Schmidt norms (Lemma 2.1), Lemma 10.11 bounds the squared expectations of the covariance defects and of the commutators of controls with earlier partitions by $O(\eta)$, the state commutators entering through Lemma 2.11(d). Its covariance estimate is stated for the rotation ζ; for a power V^h the same bound follows from its first estimate with $R = V^h Z V^{-h}$ and $S = \zeta^h Z$, since multiplication by ζ^h shifts the arcs by h, converted into the covariance defect at the cost of the state commutators, as in its proof. For two partitions of different copies, the third estimate of Lemma 10.11 bounds the first moment of the commutator aggregate by $O(\eta^{1/4})$. As $\eta \le 1$, every residual of Lemma 6.3 has first moment at most $C_G\eta^{1/4}$. There are at most $C_Q M^2$ of them, counting every bounded power, both mirrors and every earlier stage; summing their first moments before fixing the offsets, some deterministic choice of offsets has every residual at most

$$\delta \le C_G M^2 \eta^{1/4}.$$

For a single copy no two partitions of different copies occur, every residual has squared expectation $O(\eta)$, and summing the squares instead gives $\delta \le C_G M \eta^{1/2}$.

Extraction. Lemma 6.3, with these partitions and the controls $V_{i,j}^P$ in the order of the stages, gives product local isometries with vector error $\gamma \le C_Q M^3 \delta \le C_G M^5 \eta^{1/4}$ from $\Phi_D \otimes \chi$; for a single copy, $\gamma \le C_G M^4 \eta^{1/2}$. With L_i as displayed, $m_i^{L_i} \le 2^x(2+x)^{-32}$, so

$em_i^{L_i} \le (2+x)^{-32}$; and $N+1 \le x+1$, $M \le kx$ and $e(N+1)^3 \le 2^{-x}(x+1)^3 \le C(2+x)^{-32}$. So $\eta \le C_G(2+x)^{-32}$ and $\gamma \le C_G x^5(2+x)^{-8} \le C'_G(2+x)^{-3}$, which is (c).

Hashing. Discard the junk. The output vector v is a unit vector within γ of $\Phi_D \otimes \chi$, so the projection onto Φ_D times the junk misses at most γ^2 of it, and the D-dimensional output has singlet fraction at least $1-\gamma^2$. By (2.4), in either direction,

$$E_{\mathrm{D}}(\rho_{AB}) \;\ge\; (1-2\gamma^2)\log D - h(\gamma^2) \;=\; \log D - o(1),$$

since $\gamma^2 \log D \le C(2+x)^{-6}x \to 0$; all preprocessing was local on AB and nothing acted on E, so this bounds each directional rate of the original state. Finally $\log D = \sum_i r_i L_i \ge 2\alpha x - 64\alpha \log(2+x) - \sum_i r_i$ and $x = \frac{1}{2}\log(1/\varepsilon) - \log C_G$, which gives (b); the upper bound is given by the pure strategies above. □

The entropy and formation bounds of (a) are not substitutes for this proof, which never uses a decomposition of the state, a canonical purification, or an exact joint spectral law for different copies. The cuts use only the first-moment estimate of Lemma 10.11 for two partitions, which costs the margin 32 in L_i.

Remark 11.10 (The numbers without the cyclic lemma). For the coefficients alone, Lemma 6.3 is not needed. A copy of $L_{2^r,m}$ can be replaced by r commuting copies of $L_{2,m}$, which keeps the source exponent $r/\log m$ but changes the source and the game; with $q_i = 2$ throughout, the lamp z_0 of each copy, which satisfies $z_0 y z_0 = wy$, is a reflection control with $TYT \approx -Y$, and Theorem 6.2 suffices. So every coefficient of Corollary 11.11 below is realized by binary copies $L_{2,m}$ with even m. The cyclic lemma is needed for the original q-lamp games with $q \ge 4$, in particular for the one-copy games of Theorem 11.2 with $r \ge 2$.

11.5 The set of coefficients

Let $\mathcal{S}$ be the set of the coefficients α of Theorem 11.9, over all finite families (q_i, m_i) with $q_i = 2^{r_i}$, $r_i \ge 1$ and $q_i \mid m_i$. Logarithms are to base two.

Corollary 11.11 (A dense set of coefficients)**.**

$$\mathcal{S} = \Big\{ \frac{1}{2} \sum_{m \text{ even}} \frac{n_m}{\log m} : \; n_m \in \mathbb{Z}_{\ge 0},\ \textit{finitely many } n_m \ne 0,\ \textit{not all } n_m = 0 \Big\}.$$

This set is countable, closed under addition and under multiplication by positive rationals, and consists of computable reals. It contains every positive rational number and the set $\{a/(2b\log 6) : a, b \ge 1\}$*, which is dense in* $(0,\infty)$ *and consists of transcendental numbers; for example* $\frac{1}{2}\log_6 2 = 1/(2\log 6)$*. For every* $\alpha \in \mathcal{S}$ *there are a fixed Boolean constraint game and a fixed linear system game with the laws of Theorem 11.9 for* α*.*

Proof. A generator $\frac{1}{2}\log_m q$ with $q = 2^r$ dividing m has m even and equals $r \cdot \frac{1}{2\log m}$, a sum of r binary generators; conversely every $\frac{1}{2\log m}$ with m even is a generator with $q = 2$. This proves the description, and closure under addition is clear. For a positive rational a/b, $\frac{a}{b} \cdot \frac{1}{2\log m} = \frac{a}{2\log(m^b)}$ with m^b even, which gives closure under positive rational multiples; with $m = 2$ it gives every positive rational. Integer logarithms are uniformly computable, so every element of $\mathcal{S}$ is computable, and $\mathcal{S}$ is countable. The numbers $a/(2b\log 6)$ are realized by a binary copies with $m = 6^b$, and they are dense since a/b ranges over the positive rationals. Finally $\log 6$ is irrational, since $2^u = 6^v$ has no solution in positive integers; if it were algebraic, the Gelfond–Schneider theorem [65, 128] would make $2^{\log 6} = 6$ transcendental. So $\log 6$ is transcendental, and so is every nonzero rational multiple of its reciprocal. The last assertion is Theorem 11.9. □

One copy already gives a dense family. For $m = 3 \cdot 2^b$ and $q = 2^r$ with $1 \le r \le b$, the one-copy coefficient is $\alpha_{b,r} = r/(2(b + \log 3))$, which is transcendental by the same argument applied to $\log 3$. For $0 < t < \frac{1}{2}$ and $r_b = \lfloor 2t(b + \log 3) \rfloor$, one has $1 \le r_b \le b$ for all large b and $0 \le t - \alpha_{b,r_b} < 1/(2(b + \log 3))$. So the one-copy linear games of Theorem 11.2, with Theorem 11.9, have sharp usable-entanglement coefficients dense in $(0, \frac{1}{2})$.

We make no transcendence claim for general elements of $\mathcal{S}$, and the description gives no test of membership for a number presented in another form. Sums are realized by commuting copies of sources, not by an additivity theorem for one-way distillable entanglement; products and limits of elements of $\mathcal{S}$ are not realized, and a sequence of different games is not one fixed game. The transcendence step is classical, and irrational dimension exponents of this kind are already in Theorem 11.2; what is added here is the sharp original-state usable-entanglement law for them.

11.6 Which coefficients can occur

In this subsection a game is a finite game with finite question and answer sets, a winning predicate and a question distribution. Let $D_G(\varepsilon)$ be the least d such that some strategy on $\mathbb{C}^d \otimes \mathbb{C}^d$ has loss at most ε. Let $\mathcal{A}_{\log}$ be the set of finite $c \ge 0$ for which some game of quantum value one has $\lim_{\varepsilon \downarrow 0} d_G^{\rightarrow}(\varepsilon)/\log(1/\varepsilon) = c$, and $\mathcal{A}_{\dim}$ the same with $\log D_G(\varepsilon)$ in place of $d_G^{\rightarrow}(\varepsilon)$. A real is *computable relative to* an oracle X if it has a Cauchy name computable from X, and *left-c.e.* (*right-c.e.*) relative to X if the rationals below (above) it form a set that is computably enumerable relative to X. We write $0^{(n)}$ for the n-th Turing jump, and call the reals computable relative to $0^{(n-1)}$ the Δ^0_n reals.

Input convention. A game is *effectively algebraic* if each question probability $\pi(x)$ is a real algebraic number and is given by a polynomial p_x with integer coefficients together with a rational isolating interval (a_x, b_x), in which $\pi(x)$ is the only real root of p_x; any equivalent effective way of selecting a real root would serve as well. The algorithms of Proposition 11.13 and Theorem 11.14 take a game in this form.

Lemma 11.12 (Reweighting)**.** *Two games with the same questions, answers and predicate, whose question distributions have the same support, have the same logarithmic coefficients, dimension exponents and power-law exponents, whenever these exist.*

Proof. The losses are $\ell_G = \sum_x \pi_G(x)\ell_x$ and $\ell_H = \sum_x \pi_H(x)\ell_x$ with the same conditional losses $\ell_x \ge 0$, so $a\ell_H \le \ell_G \le b\ell_H$ with $a = \min_x \pi_G(x)/\pi_H(x) > 0$ and $b = \max_x \pi_G(x)/\pi_H(x)$ over the common support. Hence $d_H^{\rightarrow}(\varepsilon/a) \le d_G^{\rightarrow}(\varepsilon) \le d_H^{\rightarrow}(\varepsilon/b)$, and likewise for D. A fixed rescaling of ε changes $\log(1/\varepsilon)$ by a constant and leaves every limit of the three kinds unchanged. □

The same reweighting comparison holds for the pure original-state demand p_G: $p_H(\varepsilon/a) \le p_G(\varepsilon) \le p_H(\varepsilon/b)$. The conditional losses are unchanged on restricting to pure strategies, so the logarithmic coefficient of p_G is unchanged as well. This is the pure-state use of Lemma 11.12 below.

Replacing positive real weights by the uniform distribution on their support, the sets $\mathcal{A}_{\log}$ and $\mathcal{A}_{\dim}$ are countable. By Theorem 11.9, and with zero from a game that has a perfect product strategy, $\{0\} \cup \mathcal{S}$ lies in both.

Proposition 11.13 (Dimension exponents)**.** *For an effectively algebraic game of quantum value one, given in the input convention above, $D_G(\delta)$ is computable from G and every rational $\delta > 0$, and it equals the least Schmidt rank of a pure strategy of loss at most δ. Consequently every element of $\mathcal{A}_{\dim}$ is the limit of a computable sequence of rationals, hence, by Shoenfield's limit lemma [130], computable relative to $0'$: a Δ^0_2 real.*

Proof. For fixed d, the existence of a unit vector in $\mathbb{C}^d \otimes \mathbb{C}^d$ and of POVMs on $\mathbb{C}^d$ for the finitely many questions with loss at most δ is an existential sentence over the real closed field of real algebraic numbers, after splitting complex coordinates into real and imaginary parts and expressing positivity by principal minors. Each question probability enters as a variable π_x constrained by $p_x(\pi_x) = 0$ and $a_x < \pi_x < b_x$, which by the input convention determines it, so the sentence has integer coefficients, and quantifier elimination decides it [129, 35]. A mixed state does not do better: for fixed measurements the winning probability is linear in the state, so an eigenvector of the game operator does at least as well, and a pure strategy compresses to its Schmidt supports. Searching $d = 1, 2, \ldots$ terminates because the value is one. Approximate $\log D_G(2^{-n})/n$ by a rational r_n to error at most 2^{-n}; if the exponent exists, the computable sequence (r_n) converges to it. Shoenfield's limit lemma [130] says, relative to any oracle X, that a function is computable from X' exactly when it is the limit of a function computable from X; we use it with X computable, in the form for reals: the limit of a computable sequence of rationals is computable relative to $0'$. That direction is also immediate here, since $0'$ decides, for given k and n, the Π^0_1 statement that $|r_m - r_{m'}| \le 2^{-k}$ for all $m, m' \ge n$, and so finds a Cauchy name. By Lemma 11.12 every element of $\mathcal{A}_{\dim}$ is the exponent of an effectively algebraic game. □

No computable rate of convergence is asserted, and the value-one promise is assumed, not decided. Nothing in this proof computes a distillable entanglement, which needs a separate argument.

Theorem 11.14 (One-way coefficients)**.** *Let G be an effectively algebraic game of quantum value one, given in the input convention above. For rational $\delta > 0$, $d^{\rightarrow}_G(\delta)$ is right-c.e. relative to $0'$, uniformly in δ. If $c = \lim_{\varepsilon \downarrow 0} d^{\rightarrow}_G(\varepsilon)/\log(1/\varepsilon)$ exists and is finite, then c is left-c.e. relative to $0''$, that is, the rationals below c form a Σ^0_3 set; in particular c is computable relative to $0'''$, a Δ^0_4 real. The same holds for $d^{\leftarrow}_G$, and for the exponent β of a power law $d^{\rightarrow}_G(\varepsilon) = \varepsilon^{-\beta + o(1)}$. By Lemma 11.12, every element of $\mathcal{A}_{\log}$ is left-c.e. relative to $0''$.*

Proof. The formula. By the theorem of Devetak and Winter [51, Theorem 13 of the arXiv version], $E_{\mathrm{D}}^{A\to B}(\rho) = \lim_n \frac{1}{n} D^{(1)}(\rho^{\otimes n})$, where, for a state σ on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, $D^{(1)}(\sigma)$ is the maximum over instruments of Alice of the average coherent information $\sum_\ell p_\ell I_c(A\rangle B)_{\sigma_\ell}$, and the instruments may be taken with at most d_A^2 outcomes and one Kraus operator each. With unnormalized branches $\sigma_\ell = (A_\ell \otimes I)\sigma(A_\ell \otimes I)^*$ and $H(C) = -\operatorname{Tr} C \log C$ for positive C,

$$D^{(1)}(\sigma) = \max\Big\{ \sum_\ell \big[H(\operatorname{Tr}_A \sigma_\ell) - H(\sigma_\ell) \big] \,:\, \sum_\ell A_\ell^* A_\ell = I,\ \ell \le d_A^2 \Big\},$$

the terms $p_\ell \log p_\ell$ of the two entropies cancelling. Product instruments make $D^{(1)}$ superadditive under tensor products, so the limit is a supremum, and $D^{(1)} \ge 0$, since a rank-one measurement on Alice leaves product branches. So $E_{\mathrm{D}}^{A\to B} = \sup_N f_N$ with $f_N(\rho) = \max_{n \le N} \frac{1}{n} D^{(1)}(\rho^{\otimes n})$, an increasing sequence of continuous functions.

Computability of f_N. The objective is uniformly continuous on its compact domain with a computable modulus: the eigenvalues of a Hermitian matrix are 1-Lipschitz in the operator norm, and $t \mapsto -t \log t$ is uniformly continuous on $[0, 1]$ with an explicit modulus. The domain of the instruments, and the set K_d below, are compact sets defined by polynomial equations and inequalities; with the question probabilities entered as constrained variables, as in the proof of Proposition 11.13, the defining formulas have integer coefficients. Quantifier elimination decides for every rational box whether it meets such a set, so nested boxes give, for every point of a rational grid that lies near the set, a computable point of the set near it; evaluating a function with a computable modulus at these finitely many points gives its maximum or minimum over the set to any prescribed precision. So f_N is computable, uniformly in N and in the dimensions.

Fixed dimension. Let K_d be the set of strategies with a state on $\mathbb{C}^d \otimes \mathbb{C}^d$, POVMs on $\mathbb{C}^d$ and loss at most δ. Whether it is empty is decidable, and padding shows that it is nonempty for all $d \geq d_0$, where d_0 is the least d with $K_d \neq \emptyset$, which exists because the value is one and is found by search. Put $a_{d,N} = \min_{K_d} f_N$, computable uniformly in d, N, and $b_d = \min_{K_d} E_{\mathrm{D}}^{A\to B}$, a minimum of a lower semicontinuous function on a compact set. Then $b_d = \sup_N a_{d,N}$: clearly $a_{d,N} \leq b_d$, and if $t < b_d$, the compact sets $\{z \in K_d : f_N(z) \leq t\}$ decrease in N and have empty intersection, so one of them is empty and $a_{d,N} > t$. So b_d is left-c.e., uniformly in d, and as $0 \leq b_d \leq \log d$ the oracle $0'$, which decides whether $b_d > r$ for each rational r, computes it.

The demand. Every finite-dimensional strategy fits, after padding, in some K_d with $d \geq d_0$, so $d_G^{\rightarrow}(\delta) = \inf_{d \geq d_0} b_d$. Enumerating rational upper bounds for all b_d with $0'$ shows that it is right-c.e. relative to $0'$, uniformly in δ.

The coefficient. Put $c_m = \inf_{n \geq m} d_G^{\rightarrow}(2^{-n})/n$; these are right-c.e. relative to $0'$, uniformly in m. If c exists, it is the limit along $\varepsilon = 2^{-n}$, and then $c = \liminf_n d_G^{\rightarrow}(2^{-n})/n = \sup_m c_m$. A rational r is below c exactly when there are m and a rational $r' > r$ with $c_m \geq r'$; the condition $c_m \geq r'$ is the complement of a set c.e. relative to $0'$, so it is Π_2^0, and the set of such r is Σ_3^0, that is, c.e. relative to $0''$. The oracle $0'''$ decides membership in it and so computes c. For a power law, $\beta = \lim_n \log(1 + d_G^{\rightarrow}(2^{-n}))/n$, and since $t \mapsto \log(1+t)/n$ is increasing and computable the same argument applies. For $d_G^{\leftarrow}$ exchange the parties. □

Thus $\{0\} \cup \mathcal{S} \subseteq \mathcal{A}_{\log}$, and the matched laws of [57, Theorem 16.6 and Corollary 16.7] give the stronger inclusion $\Delta_2^0 \cap (0, \infty) \subseteq \mathcal{A}_{\log}$. Every element of $\mathcal{A}_{\log}$ is left-c.e. relative to $0''$. These two bounds do not classify the full usable spectrum. Dimension and one-way demand differ here, through the regularization, the minimization over strategies and the unbounded local dimension. The promise itself has a known place in the arithmetical hierarchy: Mousavi, Nezhadi and Yuen show that deciding whether a game has quantum value one is Π_2^0-complete [105, Theorem 1.1]. Here value one is assumed, not decided, and the objects we place in the hierarchy are the dimension exponents and one-way coefficients of games that have it. A zero coefficient for a game of Theorem A without a finite-dimensional perfect strategy would be a counterexample to the floor of Problem 18.2.

11.6.1 Pure original states and the limits of rounding

Throughout this subsection $L = L(\varepsilon) = \log_2(1/\varepsilon)$, $d_G = d_G^{\rightarrow}$ is the demand on the original mixed state, and p_G is the same infimum restricted to pure original states. Games are finite and have quantum value one whenever an asymptotic coefficient is discussed. All local dimensions are finite but unrestricted. A positive question distribution can be reweighted to uniform support by Lemma 11.12 before an effective-computability argument. The operational inputs are the pure-state entropy law [12] and Devetak–Winter hashing and regularization [51, Theorems 10 and 13].

For the extraction criterion below, the notation of [57, Lemma 12.1] is as follows. The integer m is the number of extraction stages. For each party $Y \in \{A, B\}$, $(P_{j,\perp}^Y, P_{j,0}^Y, P_{j,1}^Y)$ is the exact three-outcome PVM at stage j, U_j^B is Bob's reflection and U_j^A Alice's unitary control. Write $F_j = F_j^B = P_{j,0}^B + P_{j,1}^B$ for Bob's active projection. On a purification ψ of the original state, a column $(T_\ell)_\ell$ has error $(\sum_\ell \|T_\ell \psi\|^2)^{1/2}$. The scalar η_{ext} bounds the projection and control mirrors, the earlier-stage projection/control commutator columns and the active-swap columns specified in that lemma. Its accumulated comparison error is $\kappa_{\mathrm{ext}} = 7m^3\eta_{\mathrm{ext}}$; it is separate from the pairwise Pauli error of Section 6.7. The correction function is $g(t) = (1+t)h(t/(1+t))$, with h the binary entropy of Section 2. These symbols refer to the cited extraction interface, not to additional game variables.

Proposition 11.15 (Matched extraction)**.** *Suppose there is a fixed $C > 0$ such that every loss-ε strategy satisfies [57, Lemma 12.1] on its original state, with bounds uniform over these strategies,*

$$\sum_j \operatorname{Tr}(\rho_B F_j) \geq \log D_G(C\varepsilon) - r(\varepsilon), \qquad \tfrac{1}{2}m(m-1)\eta_{\mathrm{ext}} + 6m\kappa_{\mathrm{ext}} + 2g(\kappa_{\mathrm{ext}}) \leq s(\varepsilon),$$

where $\kappa_{\mathrm{ext}} = 7m^3\eta_{\mathrm{ext}} \leq 1/2$, $g(t) = (1+t)h(t/(1+t))$, and $r, s = o(L)$. If d_G/L has a finite limit c, then $\log D_G/L$ has the same limit and c is Δ_2^0.

Proof. The cited lemma and its actual-state bridge give $d_G(\varepsilon) \geq \log D_G(C\varepsilon) - o(L)$; the reverse comparison $d_G(\varepsilon) \leq \log D_G(\varepsilon)$ uses only the local-dimension bound on entanglement. Substitute ε/C into the first inequality to squeeze $\log D_G(\varepsilon)$ between $d_G(\varepsilon)$ and $d_G(\varepsilon/C)+o(L)$. The limits agree. Apply Proposition 11.13 after reweighting if necessary. Uniformity of the little-o errors and a fixed C are essential; no effective rate of convergence is required. The extraction hypotheses are not asserted for an arbitrary game. □

The lower inclusion in Theorem 11.16, realizing every positive Δ_2^0 real, is imported from [57, Theorem 1.1], *The uncomputable price of precision* (10.5281/zenodo.23736188). The upper-class inclusion is proved here.

Theorem 11.16 (Pure original-state coefficients)**.** *Every finite limit $p_G(\varepsilon)/L(\varepsilon)$ is Δ_2^0. Together with [57, Theorem 16.6 and Corollary 16.7], the positive spectrum of p_G is exactly the positive Δ_2^0 reals.*

The entropy classification of Hochman and Meyerovitch [70] and Meyerovitch's study of entropy dimensions of multidimensional shifts of finite type [100] are precedents for relating asymptotic invariants to arithmetical classes. The invariant here is a fixed-game precision cost, and its upper-class proof uses finite-dimensional entropy optimization.

Proof. For pure states, the one-way rate equals the marginal entropy. Fix rational question weights after reweighting. Set $\varepsilon_n = 2^{-n}$, $\tau_n = 4^{-n}$ and $R_n = 2^{n^2 4^n}$. In local dimension R_n, the set K_n of pure strategies losing at most $2\varepsilon_n$ is compact semialgebraic. Its emptiness is decidable. If nonempty, its minimum entropy m_n is computable: uniformly approximate $\eta(x) = -x\log x$ on $[0,1]$ by rational polynomials to error $2^{-n-2}/R_n$ and minimize $\operatorname{Tr} P(\rho_B)$ by real quantifier elimination. The logarithm is not assumed semialgebraic. Choose a rational q_n within 2^{-n} of m_n, or $q_n = 0$ when K_n is empty.

Suppose $p_G(\varepsilon_n)/n \to c < \infty$. Eventually an arbitrarily near-optimal pure strategy has Schmidt entropy $H < n^2$. For its decreasing probabilities λ_i, the tail T beyond R_n satisfies $T\log R_n \leq H$, hence $T \leq \tau_n$. Truncating and normalizing has half trace distance $\sqrt{T} \leq \varepsilon_n$. Compressing the original POVMs to the retained Schmidt supports preserves valid measurements; the game loss increases by at most ε_n. The truncated entropy is

$$\frac{-\sum_{i\leq R_n}\lambda_i\log\lambda_i}{1-T} + \log(1-T) \leq H/(1-\tau_n).$$

Consequently K_n is eventually nonempty and

$$p_G(2\varepsilon_n) \leq m_n \leq p_G(\varepsilon_n)/(1-\tau_n).$$

Thus q_n/n is a computable rational sequence converging to c. A $0'$ oracle finds a Cauchy tail, proving the upper class.

Conversely, [57, Theorem 16.6 and Corollary 16.7] give a rational BCS game with d_G and $\log D_G$ having any prescribed positive Δ_2^0 coefficient. The sandwich $d_G \leq p_G \leq \log D_G$ proves the assertion. □

Proposition 11.17 (Locally flagged mixtures)**.** *Let t_G be the original-state demand restricted to finite states*

$$\rho = \sum_i p_i |ii\rangle\langle ii| \otimes |\psi_i\rangle\langle\psi_i|,$$

where both flags are retained physical registers and measurements may depend on them. Then, at every feasible error,

$$E_{\mathrm{D}}^{A\to B}(\rho) = \sum_i p_i S(\psi_{i,B}), \qquad t_G \le p_G \le t_G + 1.$$

In particular their finite logarithmic coefficients coincide, and their positive spectra are both the positive Δ_2^0 reals.

Proof. The flagged state's coherent information is the displayed mean entropy, so hashing gives the lower bound. For the reverse, its flagged pure decomposition has mean entanglement $\sum_i p_i S(\psi_{i,B})$; the formation upper bound on distillation gives equality [14, 51]. Equivalently, on any tensor power and any one-Kraus-per-outcome Alice instrument, Bob's orthogonal input flags make the conditional coherent information the mean pure-branch entropy. Entropy concavity bounds the average output entropy for each fixed input branch sequence by the original Bob entropy; regularization gives the same upper bound.

For any finite collection of branch losses ℓ_i and entropies S_i, minimize $\sum_i q_i S_i$ under $\sum_i q_i = 1$ and $\sum_i q_i \ell_i \le \varepsilon$. An optimal vertex has at most two branches. Coherently superpose them with local orthogonal flags. Block-diagonal measurements retain the same loss, and the resulting pure entropy is $H(q) + \sum_i q_i S_i \le 1 + \sum_i p_i S_i$. Take infima. No readable label is supplied to an arbitrary unflagged state. □

The zero-diagonal step in Proposition 11.18 is classical; see Fillmore [61] on similarity and prescribed diagonals. For the traceless Hermitian matrix used here, Fourier transformation of an orthonormal eigenbasis gives the unitary construction explicitly.

Proposition 11.18 (Equal-loss components and joint mixing entropy)**.** *For every original strategy (ρ, A, B) with loss at most ε,*

$$p_G(\varepsilon) \le S(\rho_B) \le E_{\mathrm{D}}^{A\to B}(\rho) + S(\rho).$$

This changes neither the physical local spaces nor the loss of the selected pure component. Component selection is an existence proof for p_G, not a local procedure on ρ.

If near-minimizers have $S(\rho_\varepsilon) = o(L)$, then $p_G - d_G = o(L)$. The same follows from $\operatorname{rank}\rho_\varepsilon = \varepsilon^{-o(1)}$. *If a finite usable coefficient is not Δ_2^0, there are $\gamma > 0$ and $\varepsilon_j \downarrow 0$ with $p_G(\varepsilon_j) - d_G(\varepsilon_j) \ge 2\gamma L_j$. Every state at these errors with $E_{\mathrm{D}}^{A\to B} \le d_G + \gamma L_j$ then has $S(\rho) \ge \gamma L_j$ and* $\operatorname{rank}\rho \ge \varepsilon_j^{-\gamma}$.

Proof. Let Λ be the loss effect, $\ell = \operatorname{Tr}\rho\Lambda$, and $H = \sqrt{\rho}(\Lambda - \ell I)\sqrt{\rho}$ on $\operatorname{supp}\rho$. It is traceless Hermitian. Fourier-transform an orthonormal eigenbasis of H to a basis u_i with every diagonal expectation zero. Put $v_i = \sqrt{\rho}u_i$ and $q_i = \|v_i\|^2$. They decompose ρ, and every normalized v_i has loss exactly ℓ. Entropy concavity bounds their mean Bob entropy by $S(\rho_B)$, so at least one is no larger. Hashing proves the other inequality. The consequences use $S(\rho) \le \log\operatorname{rank}\rho$ and Theorem 11.16: if the same-loss gap were $o(L)$, the pure coefficient would equal the usable coefficient. □

Proposition 11.19 (Computable fixed states)**.** *For any uniformly computable family of finite-dimensional density matrices with effectively given dimensions, the strict lower cut $\{(e, q) : E_{\mathrm{D}}^{A\to B}(\rho_e) > q\}$ is c.e. Each fixed such rate is left-c.e. and therefore Δ_2^0. A computable robust reduction of a non-c.e. predicate to a strict rate threshold is impossible.*

Also $E_{\mathrm{D}}^{A\to B}(\rho^{\otimes k}) = kE_{\mathrm{D}}^{A\to B}(\rho)$. *If computable nonnegative integers* k_n *satisfy* $k_n/n \to a < \infty$ *and* $d_G(2^{-n}) = k_n E_{\mathrm{D}}^{A\to B}(\rho) + o(n)$ *for one computable fixed state, the coefficient* $aE_{\mathrm{D}}^{A\to B}(\rho)$ *is* Δ_2^0. *This particular repeated-state mechanism cannot separate the spectra.*

Proof. Devetak–Winter [51, Theorem 13] gives $E_{\mathrm{D}}^{A\to B}(\rho) = \sup_N f_N(\rho)$, where f_N maximizes $D^{(1)}(\rho^{\otimes n})/n$ over $1 \le n \le N$. Each finite-block optimization is uniformly computable on an effectively compact instrument domain, as in Theorem 11.14. Unnormalized branch entropies make zero-probability outcomes harmless. A strict inequality above a rational is witnessed by finite N and is semidecidable by rational error bounds; dovetail the witnesses. A promised positive gap is handled by its midpoint. Equality and non-strict comparisons are not thereby decided.

Superadditivity and regularization show that the subsequence of block lengths divisible by k has the same limiting rate, proving the tensor-power identity. The computable limit k_n/n and the left-c.e. fixed rate are both computable from $0'$, as is their product. This is not a theorem about every varying mixed-state construction. □

Proposition 11.20 (An abstract arithmetic obstruction)**.** *The numerical properties used in the arithmetic argument alone do not force* Δ_2^0 *coefficients. Precisely, for every* $\alpha \in (0,1)$ *left-c.e. in* $0''$, *there are computable rationals* $a(n,D,N)$, $D \ge 2^n$, *with* $0 \le a \le n \le \log D$, *increasing in* N *and* n *and decreasing in* D *on common domains, such that*

$$F(n) = \inf_{D\ge 2^n} \sup_N a(n,D,N), \qquad F(n)/n \longrightarrow \alpha.$$

F *is nondecreasing, both extrema are attained for each fixed choice of the other parameters, and* $Na(n,D,N)$ *is superadditive in* N. *This is an obstruction for the listed abstract properties, not a quantum-game realization theorem.*

Proof. Write $q < \alpha$ iff $\exists u\,\forall v\,\exists w\, R(q,u,v,w)$ with R decidable, and enumerate pairs (q_i,u_i) with $q_i \in \mathbb{Q}\cap[0,1]$. Let P_i be their $\forall v\exists w$ statement and $x_n = \max(\{0\}\cup\{q_i : i \le n, P_i\})$. Then $x_n \uparrow \alpha$. Set

$$a(n,D,N) = n\max\big(\{0\}\cup\{q_i : i\le n,\ \forall v \le D\ \exists w \le N\ R(q_i,u_i,v,w)\}\big).$$

All quantifiers are bounded. For fixed n, D every true candidate eventually passes. Taking D beyond a failed v for every one of the finitely many false candidates gives $\inf_D \sup_N a = nx_n$. The finite candidate indicators stabilize in N, proving supremum attainment; the same counterexample choice proves infimum attainment. Monotonicity gives $(N+M)a_{N+M} \ge Na_N + Ma_M$.

An abstract strategy $s_{m,D}$ with loss 2^{-m}, dimension $D \ge 2^m$ and cost $\sup_N a(m,D,N)$ has dimension demand 2^n and cost demand nx_n at precision 2^{-n}. Its bounded-dimension domains are finite and compact, with computable finite-block objectives. No embedding into physical state spaces or convexity of quantum strategies is asserted.

For α left-c.e. in $0'$, use $\exists u\forall v R(q,u,v)$ and the exact computable cost $b(n,D) = n\max(\{0\}\cup\{q_i : i\le n, \forall v\le D\ R(q_i,u_i,v)\})$. The same proof gives the coefficient without a regularization supremum. For example $\alpha = 1/4 + \sum_{e\in 0''} 4^{-(e+2)}$ has degree $0''$: a missing base-four digit leaves a tail at most one third of its weight, while a present digit contributes its whole weight, so the digits can be recovered from a Cauchy name. These are applications of arithmetical normal forms [130]. □

Proposition 11.21 (Weak rounding from matched laws)**.** *If* $d_G(\varepsilon) = cL + o(L)$ *and* $\log D_G(\varepsilon) = cL + o(L)$, c *finite, then for every fixed* $0 < a < 1$,

$$p_G(\varepsilon^a) \le d_G(\varepsilon) - (1-a)cL + o(L).$$

Proof. Bound $p_G(\varepsilon^a)$ by $\log D_G(\varepsilon^a) = acL + o(L)$ and subtract the first law. This applies to the full realizations of [57, Theorem 16.6 and Corollary 16.7], as well as the matched laws in its Section 12. No joint mixing-entropy assumption is needed. □

We say that G has *weak pure rounding* if $p_G(\varepsilon^a) \le d_G(\varepsilon) + o(L)$ for every fixed $0 < a < 1$. A sequence of strategies of loss at most ε is *efficient* if $E_{\mathrm{D}}^{A\to B}(\rho_\varepsilon) \le d_G(\varepsilon) + o(L)$. Efficiency here refers to the resource rate.

For the near-minimizer converse in Proposition 11.22, assume $d_G = O(L)$; this is the growth hypothesis used when dividing by $1 - \varepsilon^{1-a}$.

Proposition 11.22 (Formation versus directional distillation)**.** *For every loss-ε original mixed strategy and $\varepsilon < \delta < 1$,*

$$(1 - \varepsilon/\delta)p_G(\delta) \le E_{\mathrm{F}}(\rho),$$

where E_{F} is the single-copy convex roof. If $d_G/L \to c < \infty$ and $p_G(\varepsilon_n^a) \ge d_G(\varepsilon_n) + \gamma L_n$ along a sequence for fixed $0 < a < 1$ and $\gamma > 0$, every strategy sequence with rate at most $d_G(\varepsilon_n) + o(L_n)$ has

$$E_{\mathrm{F}}(\rho_n) - E_{\mathrm{D}}^{A\to B}(\rho_n) \ge (\gamma - o(1))L_n, \quad S(\rho_n) \ge (\gamma - o(1))L_n, \quad \operatorname{rank}\rho_n \ge \varepsilon_n^{-\gamma+o(1)}.$$

If near-minimizers with $E_{\mathrm{F}} - E_{\mathrm{D}}^{A\to B} = o(L)$ exist at every sufficiently small error, weak pure rounding holds for every fixed $a < 1$. Appending independent local mixed junk leaves both resources unchanged, although it adds its entropy to $S(\rho)$.

Proof. For any pure decomposition with the original measurements, mean nonnegative loss is at most ε. Markov leaves weight at least $1 - \varepsilon/\delta$ at loss at most δ, each with entropy at least $p_G(\delta)$. Infimize over decompositions [14]. Set $\delta = \varepsilon_n^a$; the discarded weight times the specified $O(L_n)$ lower rate is $o(L_n)$, proving the gap. Entropy concavity and hashing give $E_{\mathrm{F}} - E_{\mathrm{D}}^{A\to B} \le S(\rho) \le \log\operatorname{rank}\rho$. Divide by $1 - \varepsilon^{1-a}$ for the near-minimizer converse. Free local preparation and discard prove junk invariance. □

Proposition 11.23 (Exactly synchronous original states)**.** *For an exactly synchronous projective mixed strategy of loss at most ε,*

$$(1 - \varepsilon/\delta)p_G(\delta) \le E_{\mathrm{D}}^{A\to B}(\rho) \quad (\varepsilon < \delta < 1).$$

Thus a game with $d_G = O(L)$ and exactly synchronous near-minimizers up to $o(L)$ satisfies weak pure rounding. In fact, at the same error,

$$p_G(\varepsilon) \le E_{\mathrm{D}}^{A\to B}(\rho) + 1.$$

Proof. The exact-synchronous structure is the existing intertwiner theorem of Theorem 6.11 and [118]. On marginal supports $(A_{x,b} - B_{x,b})\operatorname{supp}\rho = 0$ identifies support vectors with intertwiners. Decomposing the finite-dimensional free-PVM representations into matching irreducibles places the support in a direct sum of Φ_{d_λ} tensor multiplicity spaces. Local central measurements give a common λ and leave $\Phi_{d_\lambda} \otimes \tau_\lambda$. All game measurements are block diagonal; sector losses satisfy $\sum w_\lambda \ell_\lambda \le \varepsilon$. Discard multiplicities, obtaining a locally flagged pure ensemble of rate $\sum w_\lambda \log d_\lambda$. Low-loss sectors have weight at least $1 - \varepsilon/\delta$ and each has $\log d_\lambda \ge p_G(\delta)$, proving the first bound and its consequence. The same ensemble is a valid same-mean-loss strategy of cost at most the original rate; Proposition 11.17 proves the stronger same-error bound. □

Proposition 11.24 (Spectral leakage on the physical spaces)**.** *Let ρ be a projective synchronous-game strategy of loss at most ε, and set*

$$H = \sum_{x,b} \nu_{x,b}(A_{x,b} \otimes I - I \otimes B_{x,b})^2, \quad P = 1_{\{0\}}(H), \quad q = \operatorname{Tr} \rho(I - P) < 1,$$

with positive weights. For $\varepsilon < \delta < 1$ define

$$t = \sqrt{q}, \qquad b = (\sqrt{\varepsilon} + \sqrt{q})^2/(1-q), \qquad k = 1 - b/\delta - 2t.$$

If $k > 0$, then $p_G(\delta) \le 1 + [E_{\mathrm{D}}^{A\to B}(\rho) + 2t + g(t)]/k$, with $g(t) = (1+t)h(t/(1+t))$.

More generally the same conclusion holds if σ is any exactly synchronous comparison state on the same physical A, B spaces, with possibly different projective measurements, loss at most b and half trace distance at most t from ρ, provided $k = 1 - b/\delta - 2t > 0$.

Proof. The normalized projection $\sigma = P\rho P/(1-q)$ is within half trace distance $\sqrt{q}$, by projecting a purification. For the positive game loss effect Λ, the Hilbert–Schmidt triangle inequality gives $\|\sqrt{\Lambda}P\sqrt{\rho}\|_2 \le \sqrt{\varepsilon} + \sqrt{q}$, so its loss is at most b. The projection defines the comparison state; the subset instrument below defines the one-way channel.

In the exact-synchronous decomposition, sectors of loss at most δ have total weight $s \ge 1 - b/\delta$ and dimension at least 2^p, where $p = p_G(\delta)$. Put $M = \lfloor 2^p \rfloor$. On each good sector convert Φ_d deterministically to Φ_M [109]. Explicitly, for every M-subset $S \subset \{1, \dots, d\}$ Alice uses

$$A_S = \binom{d-1}{M-1}^{-1/2} \sum_{i \in S} |\operatorname{rank}_S(i)\rangle\langle i|.$$

Every basis element belongs to $\binom{d-1}{M-1}$ subsets, so the instrument is complete. Alice announces S and Bob reindexes its support, completing his operation arbitrarily on its orthogonal complement. Use product outputs on bad sectors, retain only a binary good/bad flag and discard other systems; define local default operations on unmatched sectors. This is a one-way channel on the entire original input space.

Its ideal output on σ has coherent information $s \log M$ and local dimension $2M$. Its output on ρ is within t by contraction. Winter's conditional-entropy continuity bound [147, Lemma 2] decreases coherent information by at most $2t \log(2M) + g(t)$. Hashing and monotonicity yield

$$E_{\mathrm{D}}^{A\to B}(\rho) \ge k \log M - 2t - g(t).$$

Since $\log\lfloor 2^p \rfloor \ge p - 1$, this proves the formula; $M = 1$ is immediate. Starting with a comparison state proves the general version. □

For $\delta = \varepsilon^a$, $q = o(\varepsilon^a)$ gives $k = 1 - o(1)$ and, on an $O(L)$ efficient-state rate, $p_G(\varepsilon^a) \le E_{\mathrm{D}}^{A\to B}(\rho) + o(L)$. The comparison-state version needs only $b = o(\varepsilon^a)$ and $t = o(1)$. If Δ is the smallest positive eigenvalue of H, then $q \le \operatorname{Tr}(\rho H)/\Delta$. Thus $\operatorname{Tr}(\rho H) \le C\varepsilon$ and $\Delta/\varepsilon^{1-a} \to \infty$ suffice; a subpolynomial inverse gap suffices for every fixed $a < 1$. Set $q = 0$ if $H = 0$. Spectral-gap stability is studied in [49]; no universal kernel-leakage estimate follows from it here.

For projective strategies of a synchronous game, under the gap hypothesis of Proposition 11.22, with coefficient c and efficient original states, the exact-kernel leakages satisfy

$$\liminf_n \frac{q_n}{\varepsilon_n^a} \ge \frac{\gamma}{c + \gamma}.$$

Otherwise take a subsequence with this ratio bounded strictly below $\theta = \gamma/(c+\gamma)$. Then $q_n \to 0$, $b_n/\delta_n = q_n/\delta_n + o(1)$ and k_n is bounded strictly above $1 - \theta$ in the lower-bound

sense. The proved rate bound exceeds cL_n by a fixed positive multiple of L_n, a contradiction. If also $\operatorname{Tr}\rho_n H_n \le C\varepsilon_n$, then

$$\Delta_n \le [C(c+\gamma)/\gamma + o(1)]\varepsilon_n^{1-a}.$$

For the two-reflection example below, the reflection-mirror Hamiltonian means the unnormalized operator $\sum_{x=1,2}(R_x \otimes I - I \otimes S_x)^2$, where R_x and S_x are Alice's and Bob's reflections. Its normalization is distinct from the weighted Hamiltonian in Proposition 11.24.

Proposition 11.25 (Small mirror errors need no exact kernel)**.** *Alice's qubit reflections Z, X and Bob's $Z, \cos\theta X + \sin\theta Z$, $0 < \theta < \pi/2$, have no common exact synchronous vector, but their reflection-mirror Hamiltonian has Φ_2 expectation $2(1-\cos\theta)$, tending to zero.*

Proof. A kernel vector vectorizes to D with $ZD = DZ$, hence $D = \operatorname{diag}(u, v)$. The second intertwiner identity's diagonal entries force $u\sin\theta = v\sin\theta = 0$, so $D = 0$. Expanding the squared X-mirror difference on Φ_2 gives the expectation. The example rules out exact-kernel inference with fixed measurements; weak rounding remains open. □

The state in Proposition 11.26 is the Choi state of the 50% erasure channel. Its zero one-way distillability is the classical erasure-channel obstruction of Bennett, DiVincenzo and Smolin [13]; the symmetric-extension method is also treated in [110, 104].

Proposition 11.26 (The half-erasure obstruction)**.** *On $A = \mathbb{C}^d$ and $B = \mathbb{C}^d \oplus \operatorname{span}\{e\}$, let $\rho_d = (\Phi_d + I_A/d \otimes |e\rangle\langle e|)/2$. Then*

$$E_{\mathrm{D}}^{A\to B}(\rho_d) = 0, \quad E_{\mathrm{F}}(\rho_d) = \tfrac{1}{2}\log d, \quad S(\rho_d) = S(\rho_{d,B}) = 1 + \tfrac{1}{2}\log d, \quad S(\rho_{d,A}) = \log d.$$

For any fixed game with classical value $\omega_c < 1$, every measurement choice on ρ_d loses at least $(1-\omega_c)/2$. If $\omega_c = 1$, a perfect deterministic product strategy makes both demands zero. Thus this family obstructs state-only arguments but is not a fixed-game separating family.

Proof. The symmetric extension $(\Phi_{AB} \otimes e_{B'} + \Phi_{AB'} \otimes e_B)/2$ is preserved by one-way protocols after broadcasting Alice's transcript to the second Bob; symmetric-extension obstructions are standard [110, 104]. For a K-dimensional target the two singlet projections P, Q satisfy $PQP = K^{-2}P$ and $\|P+Q\| = 1 + 1/K$. Hence the extension bounds singlet fraction by $(1+1/K)/2$, which cannot approach one for a nontrivial distilled target. This proves zero directional rate, not merely nonpositive unprocessed coherent information.

The displayed pure/product ensemble gives $E_{\mathrm{F}} \le \log d/2$. Bob's erasure flag gives the reverse by monotonicity on average of pure entropy in every decomposition [14]. Orthogonal-sector entropies give the formulas. Assigning a purification entirely to Alice costs $S(B)$, and assigning it entirely to Bob costs $S(A)$, both logarithmically above the zero rate. Finally the erased sector is separable, so it cannot beat the classical value; its weight is one half. □

Proposition 11.27 (What a coefficient outside Δ_2^0 would force)**.** *If $c = \lim d_G/L$ is finite and not Δ_2^0, there are fixed $a \in (0,1)$, $\gamma > 0$ and $\varepsilon_n \to 0$ with $p_G(\varepsilon_n^a) \ge d_G(\varepsilon_n) + \gamma L_n$. Therefore all efficient original states along that sequence have the formation/distillation gap of Proposition 11.22 and, for synchronous strategies, the leakage obstruction of Proposition 11.24.*

Proof. If for every fixed $a < 1$ the limsup of $(p_G(\varepsilon^a) - d_G(\varepsilon))/L$ were nonpositive, rescaling $\delta = \varepsilon^a$ would give $\limsup p_G(\delta)/L(\delta) \le c/a$. The opposite liminf is at least c, since $d_G \le p_G$. Let $a \uparrow 1$ to obtain pure coefficient c, contradicting Theorem 11.16. A positive limsup supplies γ and a subsequence. □

These deductions leave the full mixed-state usable spectrum and unrestricted weak pure rounding open. A fixed computable state's directional rate is left-c.e.; finite-word channel reachability or failure of a universal fixed block length for quantum capacity does not prove otherwise [148, 44]. Quantum-capacity hardness and a restricted entanglement-assisted one-shot zero-error undecidability result [18, Theorems 1.1–1.2] concern different quantities. Unrestricted-LOCC continuity [127, Theorem 1.2] is not a directional-rate theorem or an input above. Exact synchrony, recovery, hashing, convex roofs and entropy continuity are credited existing tools; the assertions here are their stated applications.

12 The configuration-lamp game: usable entanglement

Our companion paper constructs an explicit binary linear system game G of quantum value one, with $m = 1687$ equations in $n = 1993$ variables and at most $s = 3$ variables per equation, and proves that every strategy that wins with probability at least $1 - \varepsilon$ has entanglement entropy $S(\rho_B) \geq c\,\varepsilon^{-1}\beta(\varepsilon)$, every mixed strategy has entanglement of formation at least $\frac{c}{4}\varepsilon^{-1}\beta(\varepsilon)$, and a maximally entangled strategy of local dimension $\exp(C\varepsilon^{-1}\log(1/\varepsilon))$ suffices, where $\beta(\varepsilon) = 2^{-4(24+\log\log(1/\varepsilon))^2} = \varepsilon^{o(1)}$ [101, Theorem A, Theorem 10.5 and Corollary 10.6]. None of these is a lower bound on the distillable entanglement of a mixed near-winner. This section proves that bound: every finite-dimensional mixed strategy of sufficiently small loss ε has one-way distillable entanglement at least $\varepsilon^{-1+o(1)}$ on its original state. Together with the upper construction, it gives $d_G^{\rightarrow}(\varepsilon) = \varepsilon^{-1+o(1)}$ for the infimum demand (Theorem 12.4).

The group behind the game is the Pauli configuration-lamp group G_{P} [101, Section 6], the Pauli analogue of the configuration-lamp group with S_3 lamps of [54, Section 3]: Pauli pairs x_v, z_v sit on the finite configurations v of the points of the three rays of Houghton's group, lamps at different configurations commute, and all centres $(x_v z_v)^2$ are one central involution J. On N points there are 2^N lamps, and the commutator of two of them has area $2^{O(\log^2 N)}$. In a model in which J is visible, the lamps are 2^N approximate Pauli pairs on the sector where J acts as -1; the companion counts them with Theorem 3.1, and we extract them with Theorem 6.4.

12.1 Inputs from the companion

We use the following results of [101]; Appendix D lists them with their numbers.

(L1) *The certificate.* A set R_{P} of 27 words in the letters p, q, t, d, x, z, each of length at most $K = 30$ (at most 22 after free reduction), containing x^2, z^2 and $(xz)^4$ [101, Definition 7.1 and Table 2]. Areas below are with respect to R_{P}, and J also denotes the word $xzxz$.

(L2) *Pair areas.* For $N \geq 1$, $n = \max(3, N)$ and each $u \subseteq \{1, \ldots, N\}$ there is a word T_u in p, q, t with $|T_u| \leq 21Nn$ such that, for $u \neq w$ and $a, b \in \{x, z\}$, the word $[T_u a T_u^{-1}, T_w b T_w^{-1}]$ has area at most Λ_N, where $14000N^2 n \leq \Lambda_N \leq 2^{(\log N + 23)^2}$ [101, Theorem 7.8 and Remark 7.10].

(L3) *Transporting J.* $T_u J T_u^{-1} J^{-1}$ has area at most $4|T_u|$, and $(T_u x T_u^{-1}\, T_u z T_u^{-1})^2 = T_u J T_u^{-1}$ freely [101, Lemma 7.11].

(L4) *The game.* $G = G(A, b)$ is the game of [101, Theorem 10.5], with the first question distribution of Definition 4.1, $m = 1687$ and $s = 3$. It is obtained by Slofstra's embedding [101, Lemma B.1] from an $E\Gamma_0$ presentation $\mathcal{P}$ with a distinguished generator w [101, Proposition 10.3]. There are words $\hat{s}$ of length at most 3 over the generators of $\mathcal{P}$, for $s \in \{p, q, t, d, x, z\}$, such that for every $r \in R_{\mathrm{P}}$ the word $\hat{r}$ obtained

by substitution, and the word $(\hat{x}\hat{z})^2 w^{-1}$, are trivial in the group of $\mathcal{P}$ with area at most $A_* = 60$ [101, Proposition 10.3(d)]. There are a homomorphism Ψ from the free group on the generators of $\mathcal{P}$ to $F(x_1, \dots, x_n, J)$, a variable x_{j_0} and a constant A_0 such that $\Psi(r)$ is trivial in $\Gamma = \Gamma(A, b)$ with area at most A_0 for every relator r of $\mathcal{P}$, and $[\Psi(w), x_{j_0}]J$ is trivial in Γ with area at most A_0 [101, Lemma B.1, (G1) and (G2)].

(L5) *The upper bound.* $\mathcal{E}^{\mathrm{me}}(G, \varepsilon) \le \exp(C\varepsilon^{-1}\log(1/\varepsilon))$ for $0 < \varepsilon \le \frac{1}{2}$ [101, Theorem 10.5(d)].

12.2 Transported lamps in an arbitrary state

The companion's sector theorem counts the transported lamps in an arbitrary state. We state it, and the estimates of its proof that the extraction below reuses.

Theorem 12.1 (The sector theorem; [101, Theorem 8.1])**.** *Let φ assign unitaries in $\mathrm{U}(D)$ to p, q, t, d, x, z, and let ρ be a density matrix on $\mathbb{C}^D$ with $\xi = \rho^{1/2}$, such that for some $e > 0$ and $0 < \Delta \le 2$:*

(i) $\|(\varphi(r) - I)\xi\|_2 \le e$ *for every* $r \in R_{\mathrm{P}}$*;*

(ii) $\|[\varphi(s), \xi]\|_2 \le e$ *for every letter* s*;*

(iii) $\|(\varphi(J) - I)\xi\|_2 \ge \Delta$.

If $e \le 2^{-1071}\Delta$*, then*

$$\log \operatorname{rank} \rho \;\ge\; c_1\, \Delta^2 e^{-2}\, 2^{-4(24+\log\log(1/e))^2}, \qquad c_1 = \frac{1}{4.1 \cdot 10^7},$$

and the same lower bound holds for $S(\rho_-)$, where ρ_- is the sector state of Lemma 2.10 for the involution $J' = \mathrm{r}(\varphi(J))$.

Fix $N \ge 1$ and put $n = \max(3, N)$, $M = 2^N$, $l = 21Nn \ge 63$ and $\Lambda = \Lambda_N$; then $l \le \Lambda/666$ by (L2). Let $A_u = \varphi(T_u x T_u^{-1})$, $B_u = \varphi(T_u z T_u^{-1})$, $X_u = \mathrm{r}(A_u)$, $Z_u = \mathrm{r}(B_u)$, $J' = \mathrm{r}(\varphi(J))$, $K_u = X_u Z_u X_u Z_u$ and $P_\pm = \frac{1}{2}(I \pm J')$. Steps 1 and 2 of the proof of [101, Theorem 8.1] use only (i) and (ii), and give

$$\|[X_u, \xi]\|_2,\ \|[Z_u, \xi]\|_2 \le \gamma := 7le, \qquad \|[J', \xi]\|_2 \le 6e,$$
$$\|(X_u - A_u)\xi\|_2,\ \|(Z_u - B_u)\xi\|_2 \le (2l+1)e,$$
$$\|[Z_u, Z_w]\xi\|_2,\ \|[X_u, Z_w]\xi\|_2,\ \|[X_u, X_w]\xi\|_2 \;\le\; 61\Lambda^2 e + 28\gamma \qquad (u \ne w), \tag{12.1}$$
$$\gamma_0 := \|(K_u - J')\xi\|_2 \le 1000 l^2 e, \qquad \|[X_u, J']\xi\|_2,\ \|[Z_u, J']\xi\|_2 \;\le\; 2\gamma_0 + 4\gamma + 12e, \tag{12.2}$$
$$\|\{X_u, Z_u\}P_-\xi\|_2 \le 5\gamma_0 + 15\gamma; \tag{12.3}$$

the argument for (12.1) there applies to every pair of letters $a, b \in \{x, z\}$, which gives all three cross types. The pair area enters squared in (12.1), because each cell of a filling of area Λ is charged its own conjugator, of length of order Λ; the linear filling charge of Theorem 13.5 would let it enter linearly, but either leaves the exponent 2 of $1/e$ unchanged.

Remark 12.2. Under the polynomial Dehn bound for Houghton's group [55, Theorems 3.2 and 3.3], the pair areas satisfy $\Lambda_N \le C' N^\kappa$ with $\kappa = 6 + \log 107$, and the factor $2^{-4(24+\log\log(1/e))^2}$ becomes a constant multiple of $(\log(1/e))^{-4\kappa}$ [101, Remarks 7.10 and 8.3]; the same replacement applies in Theorem 12.4 below.

Theorem 12.3 (Entropy and formation; [101, Theorem 10.5(b) and Corollary 10.6])**.** *There are constants $c, \varepsilon_0 > 0$ such that, for $0 < \varepsilon \le \varepsilon_0$, every strategy that wins G with probability at least $1-\varepsilon$ satisfies $\log \operatorname{rank} \rho_B \ge S(\rho_B) \ge c\,\varepsilon^{-1}\beta(\varepsilon)$, and every mixed strategy of loss at most $\varepsilon \le \varepsilon_0/2$ has $E_{\mathrm{F}}(\sigma) \ge \frac{c}{4}\varepsilon^{-1}\beta(\varepsilon)$. With (L5), $\log \mathcal{E}(G,\varepsilon) = \varepsilon^{-1+o(1)} = \log \mathcal{E}^{\mathrm{me}}(G,\varepsilon)$.*

The proof passes from a strategy to a model of Γ in Bob's state by Theorem 4.3, to a model of $\mathcal{P}$ through Ψ, in which the commutator relation $[\Psi(w), x_{j_0}]J = 1$ keeps $\varphi(\Psi(w))$ at distance about one from I, and to an assignment $\varphi'(s) = \varphi(\Psi(\hat{s}))$ of the letters of R_{P}; this assignment satisfies (i)–(iii) of Theorem 12.1 with $e = C_*\sqrt{\varepsilon}$ and $\Delta = \frac{1}{2}$ [101, proof of Theorem 10.5(b), Steps 1–4]. For the configuration lamps with S_3 lamps, at normalized-trace precision e, our arXiv paper has the same exponent 2 of $1/e$ in $\log D$ [54, Theorem A(b)].

12.3 Usable entanglement for every mixed strategy

Theorem 12.4 (The configuration-lamp game)**.** *Let G be the game of (L4), with its fixed positive question law. There are $c_G, C_G, \varepsilon_0 > 0$ such that every finite-dimensional mixed strategy of loss at most $\varepsilon \le \varepsilon_0$ satisfies*

$$E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \;\ge\; c_G\,\varepsilon^{-1}2^{-4(24+\log\log(1/\varepsilon))^2}, \qquad \text{and} \qquad d_G^{\rightarrow}(\varepsilon) \le C_G\varepsilon^{-1}\log(1/\varepsilon).$$

So $d_G^{\rightarrow}(\varepsilon) = \varepsilon^{-1+o(1)}$, and by Theorem 8.2 the same holds for $d_G^{\leftarrow}$. Neither an exact $\Theta(1/\varepsilon)$ law nor a matching subleading factor is proved.

Proof. The source on the actual state. The passage from the game to R_{P} just described is a fixed word map $s \mapsto \Psi(\hat{s})$. Apply the word-transfer identities of Lemma 2.11 to Alice's chosen row reflections, and the finite relator certificates in the proof of [101, Theorem 10.5(b), Steps 1–4]. These give the actual mirrors and inverse mirrors and the $O(\sqrt{\varepsilon})$ vector residuals on the original state; Lemma 2.11(d) gives the square-root-state commutators, while the companion's visibility argument gives $\Delta = \frac{1}{2}$. This use requires no negative-quotient identity for the source marker and no rounding of the source words. So there is a constant C such that, for $e = C\sqrt{\varepsilon}$, every signed source generator has an actual Alice mirror (the reversed word in her chosen row reflections) and inverse mirror of error at most e, every word of R_{P} has Bob residual at most e, every source letter has state commutator at most e on ρ_B, and the visibility of $J = (xz)^2$ is at least $\Delta = \frac{1}{2}$.

Bob's estimates. Fix $N \ge 1$, $M = 2^N$, $n = \max(3, N)$, $l = 21Nn$ and $\Lambda = \Lambda_N$, and define X_u^B, Z_u^B, J_B' and $P_B = \frac{1}{2}(I - J_B')$ as above. The estimates (12.1)–(12.3) are statements about state-weighted norms of Bob's operators, that is, vector norms on any purification Ψ; they hold with $\gamma = 7le$ and $\gamma_0 \le 1000l^2e$.

Alice's reflections and the filter. On Alice take the reversed mirror words and round them by the same rule $\mathrm{r}(z) = \operatorname{sgn} \Re z$, giving X_u^A, Z_u^A, J_A' and $P_A = \frac{1}{2}(I - J_A')$. A transported word has length at most $2l+1$, square defect at most $(2l+1)e$ on Bob and $3(2l+1)e$ on Alice, and rounding costs the square defect, so $\|(Q_u^A - Q_u^B)\Psi\| \le 15le$ for $Q = X, Z$; likewise $\|(J_A' - J_B')\Psi\| \le 14e$ and $\|(P_A - P_B)\Psi\| \le 7e$. With $\sqrt{p_B} \ge (\Delta - e)/2$ this gives $p_A = \|P_A\Psi\|^2 \ge \Delta^2/64 =: p_{\min}$ for $e \le \Delta/56$. Alice measures $\{P_A, I - P_A\}$, announces the outcome to Bob, and keeps the branch $\psi = P_A\Psi/\sqrt{p_A}$; the flag may be included in E, and all communication is from Alice to Bob. Transferring (12.2) to Alice, $\|[Q_u^A, J_A']\Psi\| \le 2098l^2e$; since Q_u^B commutes with P_A,

$$\|(Q_u^A - Q_u^B)\psi\| \le (15le + 1049l^2e)/\sqrt{p_A} \le 1064\,l^2e/\sqrt{p_A} =: q,$$

and Lemma 2.11(d) gives both marginal centralities after filtering, at most $\sqrt{2}q$. Every Bob pair relation R (the three cross types and the onsite anticommutator), of norm at most two,

has $\|RP_B\Psi\| = \|RP_-\xi\|_2 \le \|R\xi\|_2 + 2\|[P_-,\xi]\|_2$ or is bounded directly by (12.3), so in all cases $\|RP_B\Psi\| \le 63\Lambda^2 e$ and

$$\|R\psi\| \le \big(\|RP_B\Psi\| + 2\|(P_A - P_B)\Psi\|\big)/\sqrt{p_A} \le 64\Lambda^2 e/\sqrt{p_A}.$$

Alice's relations are within $4q$ of Bob's by transfer. Since $\Lambda^2 \ge 443556\, l^2$, every hypothesis of the uniform case of Theorem 6.4 holds on ψ with $\kappa = 128\Lambda_N^2 e/\sqrt{p_A}$.

Rate. Theorem 6.4 and the filter give $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge p_A M[1 - 2h(25\cdot 128^2 M\Lambda_N^4 e^2/p_A)]$ whenever the argument is at most $\frac{1}{2}$. If

$$2^N \Lambda_N^4 e^2 \le p_{\min}/(1600\cdot 128^2), \tag{12.4}$$

every Bell error is at most $\frac{1}{64}$ and $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge 0.76\, p_{\min} 2^N$. Choosing N maximal subject to (12.4) as in Step 5 of the proof of [101, Theorem 8.1], where the constant $1.02\cdot 10^7\Delta^{-2}$ is replaced by $1600\cdot 128^2/p_{\min}$, gives $2^N \ge c\, e^{-2} 2^{-4(24+\log\log(1/e))^2}$ for small e, and $e = C\sqrt{\varepsilon}$ gives the lower bound. The upper bound is (L5): a maximally entangled strategy of dimension $\exp(C\varepsilon^{-1}\log(1/\varepsilon))$ has directional rates equal to the logarithm of its dimension. □

The rate comes from the *entropy* of the Bell errors, not from the fidelity of the whole register: a union bound would give whole-register infidelity $O(M^2\Lambda_N^4\varepsilon)$ and only $M = \varepsilon^{-1/2+o(1)}$ pairs. The source certificate is not asserted to be a full presentation of the configuration group; only its words and the inputs (L1)–(L5) are used, together with Lemma 2.2.

13 The point-Pauli game: a polynomial window

The configuration lamps of Section 12 sit on the 2^N configurations of N points and give an almost inverse-loss demand. Placing Pauli lamps on the points themselves gives a fixed game whose demand lies between two positive powers of $1/\varepsilon$, strictly between the logarithmic and inverse scales. Whether its demand has a limiting power-law exponent, and what that exponent is, remain open.

13.1 The source and the game

The *point source* has generators a, b, s, x, z, J and the 22 relators

$$\begin{gathered} s^2,\quad (sasa^{-1})^3,\quad [s, a^2sa^{-2}],\quad aba^{-1}b^{-1}s^{-1}\ \text{(H4)},\quad asa^{-1}bs^{-1}b^{-1},\\ J^2,\ x^2,\ z^2,\quad [J,a],[J,b],[J,s],[J,x],[J,z],\quad [x,z]J^{-1},\\ [x,sa],\ [x,sb],\ [z,sa],\ [z,sb],\qquad [u,ava^{-1}]\quad (u,v\in\{x,z\}), \end{gathered}$$

with $[u,v] = uvu^{-1}v^{-1}$ and marked involution J. The first five are Johnson's presentation of Houghton's group H_3 [77], [84, Theorem 2.8], in the words of [55, (1)]; s is the transposition of the first two points of the first ray, $k = sa$ and $l = sb$ generate the stabilizer of the first point, and the last eight relators place a Pauli pair x, z at that point, commuting with its stabilizer, with all sites' pairs commuting and every onsite commutator equal to the central J. Let G_{pt} be the game of the universal compiler (Theorem 5.6) for this presentation, with the compiler's uniform-context, uniform-incidence question law. It has 6960 variables and 5713 contexts, of arity at most three, and upper amplification 12. The exported game is archived in `verify/artifacts`, where an independent checker verifies, exactly and against these 22 relators, every variable square, context commutator and forbidden assignment and the six recovery maps.

13.2 Fillings on the actual state

The lower bound needs the cost, on the actual state, of long null words with large fillings. Lemma 2.4 and Lemma 2.3(d) charge each of the A cells of a filling its own conjugator, of length up to $n + K(A+1)$ for a word of length n, and so give the quadratic bound $A(1+2n+2K(A+1))e$. The conjugators can be shared. Cut a van Kampen diagram into a tree of faces; at each face, substitute at once the whole boundary of the subdiagram that hangs from each child edge, and pay only for moving the bounded remainder of that one face. Every edge of the tree is then charged once, and the charge becomes linear in the size of the diagram (Theorem 13.5).

In this subsection $\varphi : F(S) \to \mathrm{U}(d)$ is a homomorphism, $\xi = \rho^{1/2}$ for a density matrix ρ, and $\|[\varphi(s), \xi]\|_2 \le e_c$ for every $s \in S$. Words may contain the letter 1, with $\varphi(1) = I$ and $1^{-1} = 1$, and $|v|$ counts the letters of a word v other than 1.

Lemma 13.1 (Bounded transport)**.** *For every operator Y and every word v, $\|Y\varphi(v)\xi\|_2 \le \|Y\xi\|_2 + \|Y\|_{\mathrm{op}}\,|v|\,e_c$.*

Proof. $Y\varphi(v)\xi = Y\xi\varphi(v) + Y[\varphi(v), \xi]$. The first term has norm $\|Y\xi\|_2$ because $\varphi(v)$ is unitary, and the second at most $\|Y\|_{\mathrm{op}}|v|e_c$ by Lemma 2.3(a). □

The bounded operator Y is moved past $\varphi(v)$ as a whole, so the length of the word that produced Y never enters. With actual mirrors on a purification Ψ of the original state, Lemma 2.11(b) is the same statement, with e_c replaced by the mirror error of the signed generators.

Lemma 13.2 (Substituting whole words into one face)**.** *Let $t_1, \dots, t_m$ be letters of $S^{\pm 1} \cup \{1\}$ and $q_1, \dots, q_m$ words, and put $R = t_1 \cdots t_m$, $W = q_1 \cdots q_m$ and $b_i = \|(\varphi(t_i^{-1}q_i) - I)\xi\|_2$. Then*

$$\|(\varphi(W) - I)\xi\|_2 \;\le\; \|(\varphi(R) - I)\xi\|_2 + \sum_{i:\, q_i \neq t_i} (b_i + 2(m-i)e_c).$$

Proof. Telescoping,

$$\varphi(W) - \varphi(R) = \sum_{i=1}^{m} \varphi(q_1 \cdots q_{i-1})\big(\varphi(q_i) - \varphi(t_i)\big)\varphi(t_{i+1} \cdots t_m),$$

and the terms with $q_i = t_i$ vanish. The left factors are unitary. The middle factor $Y_i = \varphi(t_i)(\varphi(t_i^{-1}q_i) - I)$ has $\|Y_i\|_{\mathrm{op}} \le 2$ and $\|Y_i\xi\|_2 = b_i$, so Lemma 13.1 bounds the i-th term, applied to ξ, by $b_i + 2(m-i)e_c$. Finally $\|(\varphi(W) - I)\xi\|_2 \le \|(\varphi(R) - I)\xi\|_2 + \|(\varphi(W) - \varphi(R))\xi\|_2$. □

The words q_i may be arbitrarily long; only the suffix of the one face R is moved.

A *disk diagram* is a finite connected plane graph whose unbounded face is bounded by a simple cycle, with every edge oriented and labelled by a letter of $S^{\pm 1} \cup \{1\}$; an edge read against its orientation gives the inverse letter. Its *faces* are the bounded faces of the graph. The boundary walk of a face, or of the whole diagram, read counterclockwise from one of its edges is a word, its *boundary word* from that edge. An edge on two faces is read in opposite directions by them, and an outer edge is read in the same direction by the diagram and by its face.

Lemma 13.3 (A tree of faces)**.** *Let D be a disk diagram with $F \ge 1$ faces, each with a boundary walk of at most K' edges, and let w be its boundary word from an outer edge g_0. Suppose that every face f has $\|(\varphi(u) - I)\xi\|_2 \le r_f$ for its boundary word u from each of its edges. Then*

$$\|(\varphi(w) - I)\xi\|_2 \;\le\; \sum_f r_f + 2K'(F-1)e_c.$$

Proof. A tree of faces. The outer boundary is a simple cycle Z. The path $Z \setminus g_0$ is acyclic, so it extends to a spanning tree T of the graph, and $g_0 \notin T$. The edges outside T are dual to a spanning tree T^* of the dual graph, whose vertices are the F faces and the unbounded face. Indeed, a cycle of such dual edges is a closed curve that crosses each corresponding primal edge once and meets the graph nowhere else, so by the Jordan curve theorem it separates the two ends of each such edge; the path in T between them would have to cross the curve, through an edge outside T, which is impossible. So these dual edges form a forest, and by Euler's formula there are $E - V + 1 = F$ of them, one fewer than the number of dual vertices; hence T^* is a spanning tree. The only outer edge outside T is g_0, so the unbounded face is a leaf of T^*, attached to the face f_0 that contains g_0. Deleting it leaves a tree T_0^* on the F faces, whose $F - 1$ edges are interior edges outside T. Such an edge lies on two distinct faces, since an edge with one face on both sides is a bridge and every bridge lies in T, and each of the two faces traverses it once.

Substituting subtrees. Root T_0^* at f_0. For $f \neq f_0$ let g_f be the edge joining f to its parent. Let $R_f = t_1 \cdots t_m$, $m \leq K'$, be the boundary word of f from g_f, and that of f_0 from g_0. Define W_f, recursively from the leaves, by replacing in R_f each letter t_i read along the edge g_c of a child c by the word q_i obtained from W_c by deleting its first letter. That first letter is read along g_c in the direction opposite to f, so it is t_i^{-1}, and $t_i^{-1} q_i = W_c$. Put $\beta_f = \|(\varphi(W_f) - I)\xi\|_2$. By Lemma 13.2, with $m - i < K'$,

$$\beta_f \;\leq\; r_f + \sum_{c \text{ child of } f} (\beta_c + 2K' e_c),$$

and induction from the leaves gives $\beta_{f_0} \leq \sum_f r_f + 2K'(F-1)e_c$, since T_0^* has $F - 1$ edges.

Free equality. Read each W_f as an edge path. The boundary walk of f from g_f is closed; replacing its traversal of g_c, from u to v say, by the remainder of the closed path W_c, which runs from u back to v, keeps it a closed path. So W_{f_0} is a closed edge path beginning with g_0. Every edge g_c of T_0^* is traversed once by its parent and once by its child, at the first letter of W_c, and both traversals are deleted; so W_{f_0} uses only edges of $T \cup \{g_0\}$ and traverses g_0 exactly once, in the direction of w. In the graph $T \cup \{g_0\}$, whose only cycle is Z, a closed edge path from the initial vertex of g_0 is determined up to homotopy by the signed number of its traversals of g_0. Hence W_{f_0} and the boundary walk w are homotopic, they become equal after cancelling backtracking pairs of edges, and their words are freely equal. So $\varphi(w) = \varphi(W_{f_0})$. □

A van Kampen diagram may be singular, with edges and vertices on no face, and a face may meet itself; the next lemma removes this at linear cost.

Lemma 13.4 (Regular diagrams at linear cost)**.** *Let R be a set of words of length at most $K \geq 1$, and let a freely reduced word $w_0 \neq 1$ of length n be freely equal to a product of A conjugates of elements of $R^{\pm 1}$. There is a disk diagram whose boundary word from a suitable edge is w_0 with letters* 1 *inserted, and whose faces are at most A* caps *and further* auxiliary *faces. The boundary word of a cap, from any of its edges, is a cyclic permutation of an element of $R^{\pm 1}$ with letters* 1 *inserted, of at most $2K$ letters; that of an auxiliary face has at most four letters and is freely trivial. There are at most $A + 5E$ faces, where E is an integer with $2E \leq KA + n$.*

Proof. By van Kampen's lemma [92, Chapter V, Section 1], as in the proof of Lemma 2.4, w_0 has a van Kampen diagram D_0 with at most A faces, the boundary walk of each reading a cyclic permutation of an element of $R^{\pm 1}$, and with outer boundary walk w_0 from a base vertex. Let E be its number of edges. Each edge has two sides, and each side lies on exactly one face boundary walk or on the outer walk, so $2E \leq KA + n$.

Thicken the 1-skeleton of D_0 in the plane: a small closed disk around each vertex and a thin band along each edge, meeting the disks of its two ends in two short arcs. The union N is a connected compact planar surface, and the bounded components of its complement are the faces of D_0, slightly shrunk, one for each. Fill each of them with a disk, the cap of that face; by the Jordan–Schoenflies theorem, N with its caps is a closed disk, bounded by the outer boundary curve of N. The two long sides of the band of an edge carry the letter of that edge, oriented alongside it, and every short band end, every arc of a vertex disk and every diagonal added below carries 1. A band then reads $t\,1\,t^{-1}\,1$, which is freely trivial from every edge. A cap meets, along its boundary, one long band side for each edge of its face and one vertex arc between consecutive ones, so it reads the relator of that face with letters 1 inserted. The outer boundary meets one long band side for each step of the outer walk of D_0, including both sides of an edge that the walk traverses twice, so, read from the band side of the first edge of w_0, it is w_0 with letters 1 inserted. The disk of a vertex of degree $\delta \geq 2$ is a polygon with 2δ sides, alternately band ends and free arcs; divide it by diagonals into $2\delta - 2$ triangles. The disk of a vertex of degree one has two sides and is kept. These faces read words in the letter 1 only. As the degrees sum to $2E$, there are at most $A + E + 4E$ faces. □

Theorem 13.5 (A linear filling charge)**.** *Let R be a finite set of words of length at most $K \geq 1$, and put $K' = \max\{2K, 4\}$. Suppose that $\|[\varphi(s), \xi]\|_2 \leq e_c$ for every $s \in S$ and $\|(\varphi(r) - I)\xi\|_2 \leq e_r$ for every $r \in R$. If a word w of length n has area at most A in R, then*

$$\|(\varphi(w) - I)\xi\|_2 \;\leq\; Ae_r + \big[(2K + 2K' + 5KK')A + 5K'n\big]e_c.$$

In particular, in a model at precision e the error is at most $C_K(A + n)e$, with $C_K = 1 + 2K + 2K' + 5K'(K + 1)$.

Proof. Let w_0 be the free reduction of w, so that $\varphi(w) = \varphi(w_0)$ and $|w_0| \leq n$; if $w_0 = 1$ there is nothing to prove. Otherwise take the disk diagram of Lemma 13.4 and apply Lemma 13.3, whose faces have at most K' edges. A cap read from any edge is, up to letters 1, a word $\alpha^{-1} r^{\pm 1} \alpha$ with $r \in R$ and $|\alpha| \leq K$. Since $\varphi(r)^{-1} - I = -\varphi(r)^{-1}(\varphi(r) - I)$, Lemma 2.3(c) and (a) give $r_f \leq e_r + 2Ke_c$ for a cap, and an auxiliary face has $r_f = 0$. With at most $A + 5E$ faces and $2E \leq KA + n$,

$$\|(\varphi(w) - I)\xi\|_2 \;\leq\; A(e_r + 2Ke_c) + 2K'(A + 5E)e_c \;\leq\; Ae_r + \big[(2K + 2K')A + 5K'(KA + n)\big]e_c.$$

In a model at precision e both e_r and e_c are at most e. □

Remark 13.6. The proof uses the state only through Lemma 13.1. With Lemma 2.11(b) in its place the same bound holds for $\|(U(w) - I)\Psi\|$ on a purification of the original mixed state, with e_c the actual mirror error of the signed generators; neither form conditions on a sector or replaces the state. When every mirror is exact, for instance on a maximally entangled state with transposed mirrors, $e_c = 0$ and the bound is the tracial estimate Ae_r. At a common precision e the separately based conjugators of Lemma 2.4 cost $A(A + n)e$, while the tree of faces costs $C_K(A + n)e$. No signed area and no cancellation between cells is used. For the five Houghton relators $K = 12$ and $C_{12} = 1633$.

Proposition 13.6a (Far-commutator area)**.** *In the five Johnson relators displayed in Section 13.1, put $\sigma_i = a^i s a^{-i}$, and let $A^*(R)$ be the largest area of $[\sigma_i^{\pm 1}, \sigma_j^{\pm 1}]$ with $0 \leq i, j \leq R$ and $|i - j| \geq 2$. There is a constant C such that, for every integer $R \geq 2$,*

$$A^*(R) \leq C R^{\log_2 \lambda}, \qquad Q(\lambda) = 0, \quad Q(t) = t^5 - 89t^4 + 824t^3 - 3416t^2 - 6672t + 19648,$$

where $\lambda \in (79, 80)$ is the largest positive root of Q. An admissible area exponent is

$$p = \log_2 \lambda = 6.3064682540119605963\ldots.$$

Proof. Appendix G.1 gives the complete six-type filling construction, with every auxiliary relator expanded into the literal five-relator presentation. The recurrence follows the filling method of [55]; the explicit matrix and certificates in the appendix establish the exponent stated here. □

13.3 The window

Theorem 13.7 (The point window)**.** *There are constants $c, C > 0$ such that, for all sufficiently small ε,*

$$c\,\varepsilon^{-\beta} \;\le\; d^{\rightarrow}_{G_{\mathrm{pt}}}(\varepsilon) \;\le\; C\varepsilon^{-1/4}, \qquad \beta = \frac{1}{2\log_2 \lambda + 1} = 0.0734595360384268649\ldots.$$

Here λ is the root in Proposition 13.6a. In particular, every original finite-dimensional mixed state of a strategy of loss at most ε has $E_{\mathrm{D}}^{A\to B}(\rho) \ge c\varepsilon^{-\beta}$. By Theorem 8.2 the same window holds for $d^{\leftarrow}_{G_{\mathrm{pt}}}$. A matching lower exponent $\frac{1}{4}$ is open (Section 18).

The lower bound uses only a fixed visibility certificate, the polynomial far-commutator areas of Houghton's group, the linear filling charge of Theorem 13.5 and the S_3 extraction of Section 6.4. The upper bound is an explicit finite model.

Lemma 13.8 (A fixed visibility certificate)**.** *Put $\gamma = s\,(asa^{-1})$, a three-cycle. There is a constant C_0 such that every model of the point source at precision e has $\|(J-I)\xi\|_2 \le C_0(\|(\gamma-I)\xi\|_2 + e)$. Consequently $\|(\gamma-I)\xi\|_2 \ge \Delta_0 > 0$ for all sufficiently small e.*

Proof. With $k = sa$, $l = sb$ and $t = asa^{-1}$, the identity $t = [k,l]$ holds in H_3: both sides exchange the second and third points of the first ray and fix all others, which can be checked on the finitely many points the words move; by Johnson's theorem it has a fixed finite derivation from the five Houghton relators. In the quotient by $\gamma = 1$: x, z commute with k, l, hence with t, hence with $s = t^{-1}$, hence with $a = s^{-1}k$; then $[x, aza^{-1}] = 1$ becomes $[x, z] = 1$, and $[x, z] = J$ gives $J = 1$. So J lies in the normal closure of γ and the relators, with a fixed derivation. Evaluating it with Lemma 2.3, every occurrence of γ costs its residual and every fixed conjugator a fixed multiple of e. Since $\|(J+I)\xi\|_2 \le e$, $\|(J-I)\xi\|_2 \ge 2 - e$. □

Proof of the lower bound. By Proposition 6.1, applied to the universal compiler, every strategy of loss ε gives Bob unitaries for the six generators and Alice opposite mirrors, with relator residuals, mirrors, inverse mirrors and state commutators at most $e = C\sqrt{\varepsilon}$, on the original state. Round s once to a reflection on each party by $\operatorname{sgn}\Re$; by Lemma 2.5 and Lemma 2.3(e), applied to the fixed relator words, the rounded tuple still has relator residuals, mirrors and state commutators $O(e)$. This changes observables, not the state. Put $p = \log_2 \lambda$ and, for $0 \le i < N$,

$$S_i = a^{3i} s a^{-3i}, \qquad T_i = a^{3i+1} s a^{-3i-1},$$

reflections on each party (Alice using the reversed mirror words). Then $(S_iT_i)^3$ is a conjugate of $(sasa^{-1})^3$ and S_iT_i a conjugate of γ, by words of length $O(N)$, so the cubic residuals and raw mirrors are $O(Ne)$ and, by Lemma 13.8, Bob's raw three-cycle visibility is at least $\Delta_0 - O(Ne)$. Every cross-copy commutator is a far commutator $[\sigma_{p'}, \sigma_{q'}]$ with $\sigma_j = a^j s a^{-j}$, $0 \le p', q' \le 3N+1$ and $|p'-q'| \ge 2$, a word of length $O(N)$ whose area in the five Houghton relators is at most $C(3N+1)^p$ by Proposition 13.6a, in the literal presentation and word orientation. Since $p > 1$, Theorem 13.5 bounds its residual by $C(N^p + N)e \le CN^p e$. The additive and the group commutator of two reflections differ by a state commutator of a word of length $O(N)$, so Bob's raw cross-commutators are at most $\eta_0 \le CN^p e$. Lemma 6.9 gives exact local S_3 representations with $\kappa \le 9\eta_0 + 144\mu_0 + 36e_{\mathrm{cub}} \le CN^p e$ and Alice visibility $\Delta \ge \Delta_0/2$ once Ne is small. Take $N = \lfloor c\,e^{-2/(2p+1)} \rfloor$ with c small: then $N\kappa^2 \le Cc^{2p+1} \le \Delta^2/38400$,

while $Ne \to 0$ because $2/(2p+1) < 1$. For sufficiently small e, the floor also gives $N \geq (c/2)e^{-2/(2p+1)} \geq 1$. Corollary 6.8 gives $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \geq \Delta^2 N/16 \geq c' e^{-2/(2p+1)} = c'' \varepsilon^{-1/(2p+1)}$, the exponent being halved because e is of order $\sqrt{\varepsilon}$. □

Charging each cell its own conjugator (Lemma 2.4) would give only $\eta_0 \leq CN^{2p}e$ and the exponent $1/(4p+1)$, with $p = \log_2 \lambda$ from Proposition 13.6a. The area exponent enters only through η_0; Section 18 records what smaller areas would give. The route itself has a limit: no worst-pair estimate $\eta_0 \leq CN^u e$ with $u < 2$ holds on all original source models, already at the scale of this proof (Theorem G.2 and Proposition G.3), so it cannot give a lower exponent beyond $\frac{1}{5}$.

13.4 A collective finite model

The upper bound needs finite tuples of the literal 22 relators with small defect and little entanglement. The models of [55, Proposition 6.5] spread the single defect of Houghton's corner models over a torus of clocks by rotating the links in two-dimensional phase planes; on an exterior-algebra quantization they give source precision $O(M^{-3/2})$ at entropy $O(M)$, and hence only $d \leq C\varepsilon^{-1/3}$. The model below does better by taking the logarithm of a *product* of swaps, once, after representing it.

Proposition 13.9 (Collective models)**.** *For every $M \geq 48$ there are unitaries for the 22-relator source on $\mathbb{C}^{D_M}$, $D_M = M^2\, 2^{4M+4}$, with $J = -I$, all relators except H4 exact, and $\|\mathrm{H4} - I\|_{\mathrm{op}} \leq \pi/A_B + \pi/A_T \leq 8\pi/M^2$, where $A_B = M(M-1)/2$ and $A_T = 1 + (M-1)(M-2)/2$. On the maximally entangled state Φ_{D_M}, with Alice's transposes, every signed generator mirror is exact, and both directional rates equal $4M + 4 + 2\log M$.*

Proof. Base data. We use the permutation data of [55, §6] in its right-action convention: $L = 4M + 4$ colour coordinates, clocks $c \in (\mathbb{Z}/M\mathbb{Z})^2$, the link permutations A_c, B_c of the corner models with the two preliminary exact moves $\eta_1 = (1\ M{+}4)$ at $(M-1, M-1)$ and $\eta_2 = (5\ M{+}5)$ at $(0, M-2)$, and $\tau = (1\ 2)$. They satisfy $A_c \tau A_c^{-1} = (2\ 3)$, $A_c(2\ 3)A_c^{-1} = (3\ 4)$ and $B_c \tau B_c^{-1} = (2\ 3)$ at every clock, and their H4 holonomy is the identity except at the two sources $q_B = (M-1, M-1)$, where it is $\prod_{t=5}^{M+3}(t\ t{+}M)$, and $q_T = (0, M-2)$, where it is $(5\ M{+}5)$ [55, proofs of Propositions 6.1 and 6.5].

One common region. Let B be the smallest bulk region R_5 of [55, proof of Proposition 6.5] and T the region with source q_T, with their labels: for every $t = 5, \dots, M+3$ the label $S_t(v)$ of R_t restricted to the vertices of B, and the label $S_T(v)$. Restriction preserves the four regional properties of [55, Lemma 6.4], because $B \subseteq R_t$ for every t: every base link outside the sources carries each $S_t(v)$, hence their product, to the next label; the regions are connected through interior edges with zero boundary potentials; base labels avoid $\{1, 2\}$ and labels at initial vertices of interior edges avoid $\{2, 3, 4\}$; and labels of different regions are disjoint. Let $P_B(v) = \prod_t S_t(v)$, as a product of disjoint transpositions, and $P_T(v) = S_T(v)$; at the sources these are the two holonomies.

Represent first, then take one logarithm. Let Π be the qubit-permutation representation on $(\mathbb{C}^2)^{\otimes L}$, and put $V_R(v) = \Pi(P_R(v))$ and $Q_R(v) = \frac{1}{2}(I - V_R(v))$ for $R \in \{B, T\}$. These Q_R are orthogonal projections with $\exp(i\pi Q_R) = V_R$ exactly, covariant along base links $(\Pi(A)Q_R(w)\Pi(A)^* = Q_R(v)$ for an edge from v to w), commuting with one another where both occur, and acting trivially on qubits $2, 3, 4$. Solve the scalar curl equations $\operatorname{curl}\varphi_R = \pi/A_R - \pi\, 1_{q_R}$ on R with zero boundary potentials by a spanning tree of the dual graph, put $U_c = \prod_R \exp(i\varphi_R(a\text{-edge at } c)Q_R(c))$ and $V_c = \prod_R \exp(i\varphi_R(b\text{-edge at } c)Q_R(c))$, and define on $\mathbb{C}^{M^2} \otimes (\mathbb{C}^2)^{\otimes L}$ the operators $a|c, v\rangle = |c - e_1, U_{c-e_1}\Pi(A_{c-e_1})v\rangle$, $b|c, v\rangle = |c - e_2, V_{c-e_2}\Pi(B_{c-e_2})v\rangle$, $s|c, v\rangle = |c, \Pi(\tau)v\rangle$, $x = X_1$, $z = Z_1$ and $J = -I$.

The relators. Since the rotations commute with $\Pi((2\ 3))$ and $\Pi((3\ 4))$, [55, Lemma 6.3] gives the four Houghton relators other than H4 exactly. The new link factors act trivially on qubits $2, 3, 4$ and the base permutations have the same first inverse images as in [55], so in the left-operator convention $ax_1a^* = x_2$, $az_1a^* = z_2$, $bx_1b^* = x_2$, $bz_1b^* = z_2$, $sx_2s^* = x_1$ and $sz_2s^* = z_1$; hence sa and sb fix both lamps and $[u, ava^{-1}] = [u_1, v_2] = 1$. Squares, centrality of the scalar J and $[x, z] = -I = J$ are exact. For H4, the computation of [55, Lemma 6.4] uses only covariance, commutation, protected supports and the endpoint identity $\exp(i\pi Q) = V$ of its generators, all of which hold here, so the holonomy at every plaquette is $\exp(i1_B(q)\pi Q_B(q)/A_B)\exp(i1_T(q)\pi Q_T(q)/A_T)$, of operator distance at most $\pi 1_B/A_B + \pi 1_T/A_T$ from I. Finally $A_B, A_T \geq M^2/4$ for $M \geq 4$.

The state. On Φ_{D_M} with Alice's transposes, $(U^T \otimes I)\Phi = (I \otimes U)\Phi$, so every mirror is exact and the state is pure with entropy $\log D_M$. □

Proof of the upper bound in Theorem 13.7. With the flat state, the collective model is a source model at precision $e_M \leq 8\pi/M^2$. The upper map of the universal compiler gives a model of the constraint system at precision Ce_M with the state $I_{12}/12 \otimes \rho$, and the boolean return gives a strategy of loss at most $K_3C^2e_M^2 = O(M^{-4})$ on a maximally entangled state of dimension $12D_M$, whose rates are $4M + 4 + 2\log M + \log 12$. Taking M of order $\varepsilon^{-1/4}$ gives $d^{\to} \leq C\varepsilon^{-1/4}$. □

Proposition 13.9a (Fixed-holonomy subdivision from the magnetic-cycle estimate)**.** *Let $A \geq 1$, $F_A \cdots F_1 = V = e^{iK}$ be finite-dimensional unitaries with $K = K^*$ and $\operatorname{spec} K \subset [-\pi, \pi]$, and let ρ be a density matrix. If $[\rho, V] = 0$, then*

$$\frac{1}{A}\sum_j \|(F_j - I)\sqrt{\rho}\|_2^2 \geq 4\operatorname{Tr}(\rho \sin^2(K/(2A))),$$

with equality for $F_j = e^{iK/A}$. Without commutation,

$$\frac{1}{A}\sum_j \|(F_j - I)\sqrt{\rho}\|_2^2 \geq A^{-2}\|(V - I)\sqrt{\rho}\|_2^2.$$

For normalized trace and $V = I - 2Q$, the infimum is exactly $4\tau(Q)\sin^2(\pi/(2A))$. For the fixed bulk holonomy and area of Proposition 13.9, this is $\Theta(M^{-4})$, so changing only its subdivision cannot improve its root-defect order.

Proof. The complete proof is in Appendix G.2. The scalar input is the magnetic-cycle spectrum [40, Lemma 2.9] and its equal-distribution minimization [21, Proposition 4.6(b)]; the arbitrary-state estimate is elementary telescoping. The application fixes the holonomy, representation, visibility and factor-count relation, and does not exclude changing them. □

Remark 13.10. The actual tracial H4 error of the collective model is $\Theta(M^{-4})$ in squared normalized Hilbert–Schmidt norm, so the root precision is genuinely $\Theta(M^{-2})$. The difference from the exterior-algebra models is specific: those take a sum of the quadratic logarithms of $O(M)$ commuting swaps, whose tracial fluctuations are of size $\sqrt{M}$, whereas here the logarithm of their product is one bounded projection. The new links do not factor as a native link times a control, so the native class law below does not apply to them. An exterior-algebra specialization of the same construction also works, at dimension $M^2 2^{8M+8}$.

13.5 A conditional law on the calibrated native class

The exterior-algebra models of [55, Proposition 6.5], with $k = 2$ flavours and with the Pauli lamps $x = i\gamma_{1,1}\gamma_{1,2}$, $z = i\gamma_{1,2}\gamma_{1,3}$ built from the Majorana operators of the first site and $J = -I$, satisfy the 22 relators with only H4 inexact, $\|\mathrm{H4} - I\|_{2,D}^2 \le 14.26M^{-3}$, at dimension $M^2 2^{8M+8}$. Call such a model, for one $M \ge 48$, a *native model* $\mathcal{N}_M$. On the class of Bob tuples equal to one native model, the demand has a matching law, but the proof of that law rests on inputs that have not been supplied self-containedly.

Theorem 13.11 (Native class law; conditional)**.** *Restrict Bob's tuple to one native model $\mathcal{N}_M$, $M \ge 48$, up to a common unitary gauge and arbitrary finite spectators, and let the original ABE state be arbitrarily correlated. Require actual Alice signed-generator and inverse mirrors at most e, all 22 actual vector relator residuals at most e, and the negative centre; do not assume a flat density, stationarity, clock diagonality, a pure input, flavour factorization or an independent environment.* Conditional on the inputs listed below*, the minimum original $E_{\mathrm{D}}^{A\to B}$ on this class is $\Theta(e^{-2/3})$, and on the game G_{pt}, restricted to the calibrated native class (strategies whose actual native source contract is at most a fixed $C\sqrt{\varepsilon}$), the minimum demand is $\Theta(\varepsilon^{-1/3})$. This is not the unrestricted optimum: the collective models of Proposition 13.9 lie outside the class, and the unrestricted upper bound $\varepsilon^{-1/4}$ of Theorem 13.7 is below this class law.*

The conditional inputs. The lower bound is the aggregate extraction of Theorem 6.6 applied to transported point lamps, $b_N = 9E_B/N + 132M_2$, together with four native estimates: the coercivity $M^2e^2 + M^3\|(\mathrm{H4} - I)\Psi\|^2 \ge c$; a sparse bad-branch probability $p_{\mathrm{bad}} \le 2N/M + 4Ne$; the aggregate bound $E_B \le 768\, N p_{\mathrm{bad}}$; and hence $b_N \le 13824N/M + 27648Ne + 396N^3e^2$, which with $N = c\min\{M, e^{-2/3}\}$ gives $E_{\mathrm{D}} \ge c'e^{-2/3}$. The coercivity and the bad-branch estimates rest on the appendix "Sharp Houghton cost with moderate flavour growth" of the authors' manuscript on immediate delivery [56], whose proofs are not reproduced here, and the bound on E_B has a written derivation that is not included here; the theorem becomes unconditional when self-contained proofs of these are supplied. The aggregate theorem itself (Theorem 6.6) is proved. The upper bound uses the literal free-unitary lift $I_{12} \otimes \mathcal{N}_M$ in the compiled game, not an assignment correct only modulo the source relations: the recovery polynomial evaluates to $I_{12} \otimes s$ on every unitary source tuple. The source witness has flat entropy $8M + 8 + 2\log M$; with the amplified flat state the compiled game witness is maximally entangled of dimension $12M^2\, 2^{8M+8}$, belongs to the native class, and has $e \asymp M^{-3/2}$ and $E_{\mathrm{D}} = 8M + 8 + 2\log M + \log 12$.

Theorem 13.12 (Enlarged native class; conditional)**.** *Keep exact native clocks and the first two protected flavours, but allow clock-dependent links $L_a^0(c) \otimes U_a(c)$ and $L_b^0(c) \otimes U_b(c)$ and a clock-diagonal $s_0 \otimes S(c)$, with lamps $x_0, z_0 \otimes I$, negative J, finite noncentral controls, arbitrary correlations with Alice and the environment, common gauges, and finite coherent direct sums of different M. Require all full actual relator residuals, control defects included, and all actual mirrors, at most e. Links must factor between protected and control factors as stated; arbitrary entangling link operators are not included. Conditional on the same inputs as Theorem 13.11, the class law is again $\Theta(e^{-2/3})$, and $\Theta(\varepsilon^{-1/3})$ only for its calibrated native game restriction.*

The finite local identity behind the enlarged class is checked exactly: for the generator H of a native link, averaged over the local Pauli group of the two protected flavours, $\mathbb{E}_P(H^2) - \mathbb{E}_P(H)^2 = I/8 + H_{(+,-)}^2 \ge I/8$ on every multiplicity sector. Its sector estimates give $\sum_M p_M M^{-3} \le Ce^2$, hence $\sum_M p_M M^{-1} \le Ce^{2/3}$ by Jensen's inequality, and after clipping small sectors $b_N \le C_1 Ne^{2/3} + C_2 Ne + C_3 N^3 e^2$. The conditional status concerns the operational law, not this identity.

14 Exact sources: the price of an address

A finite game sees its source only through approximate relations. It is instructive to ask what the same groups cost when Bob's relations are exact. We work under the exact-source contract (EX) of Section 2.8, with Alice's mirror error e. The demand is then governed by one quantity: how many independent Pauli qubits can be *addressed* by short words. Linear, polynomial and exponential address growth give demands $\Theta(1/e)$, $\Theta_k(e^{-k})$ and $2^{\Theta(1/e)}$, in both physical directions. Exact relations and approximate ones can differ by an exponential in this sense (Section 15).

14.1 The address-growth principle

Theorem 14.1 (Address growth)**.** *Suppose an exact negative source representation contains m independent exact Pauli pairs X_i, Z_i whose* individual *generators are represented by source words of length at most r, and that Alice's actual signed-generator mirrors have root error at most e. Then*

$$E_{\mathrm{D}}^{B\to A}(\rho_{AB}) \geq m[1 - 2h(r^2e^2/4)], \qquad E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \geq m[1 - 2h(r^2e^2/2)],$$

in the monotone range of the entropy argument. Choosing re a small constant gives $\Omega(m(c/e))$ in both directions, where $m(r)$ is the number of addressable sites.

Proof. The m exact Pauli pairs generate $M_{2^m}(\mathbb{C})$, so Bob's space is $\mathbb{C}^{2^m} \otimes K$ with X_i, Z_i the standard Paulis on the first factor; this is a local unitary identification of the original space, not a compression. The reversed Alice word of each individual generator is an actual unitary predictor with error at most re, by Lemma 2.11(a). Apply Theorem 6.10(c) with $p \leq r^2e^2/4$, using $2p(1-p) \leq 2p$ in the forward direction. □

The full Pauli-string length need not equal the individual addressing length, and no joint Alice measurement is used: individual predictors need not commute. Addressing supplies only a lower bound; a matching law needs an upper construction in an exact representation, which the following subsections give. The uncertainty relation with memory [16] and hashing [51] are the standard inputs.

14.2 The point source

Let Ω be the three rays of H_3 and P_Ω the group generated by a central involution J and involutions x_p, z_p, $p \in \Omega$, with $[x_p, z_p] = J$ and all generators at distinct sites commuting. A word supported on a finite set S has normal form $J^\epsilon \prod_{p\in S} x_p^{a_p} z_p^{b_p}$, so P_Ω is locally finite.

Proposition 14.2 (Structure of the point source)**.** *The group presented by the 22 relators of Section 13.1 is $P_\Omega \rtimes H_3$. It is an extension of $\mathbb{Z}^2$ by the locally finite group $P_\Omega \rtimes \mathrm{FSym}(\Omega)$, hence elementary amenable. Every finite-dimensional representation sends J to I.*

Proof. In the presentation, $\langle sa, sb\rangle$ is the full stabilizer of the first point, so the transported pairs $x_y = g_y x g_y^{-1}$, $z_y = g_y z g_y^{-1}$, with g_y taking the first point to y, are well defined; the four pair prototypes and the two-transitivity of H_3 give all cross-site commutations, and the onsite and central relations give $[x_y, z_y] = J$. Conversely $P_\Omega \rtimes H_3$ satisfies the 22 relators, and the natural maps are mutually inverse. Finitely many generators of $P_\Omega \rtimes \mathrm{FSym}(\Omega)$ involve finitely many sites and a finite permutation group, so the group is locally finite, and the quotient by it is $H_3/\mathrm{FSym} = \mathbb{Z}^2$. In a finite-dimensional representation the $J = -1$ eigenspace is invariant and carries arbitrarily many independent Pauli pairs, so it is zero. □

Theorem 14.3 (Exact point source)**.** *Under (EX) for the point source, for every N there are product local isometries extracting $\mathbb{C}^{2^N}\otimes\mathbb{C}^{2^N}$ registers within vector error $4Ne$ of a maximally entangled state times a residual purification, so with Bell infidelity at most $16N^2e^2$, and*

$$E_{\mathrm{D}}^{A\to B},\ E_{\mathrm{D}}^{B\to A}\ \geq\ N(1-32N^2e^2)-h(16N^2e^2)\qquad(16N^2e^2<\tfrac{1}{2}).$$

Exact induced representations with finite-rank pure states of entropy $O(1/e)$ exist, so both directional infima are $\Theta(1/e)$. For a normal density D in an exact negative representation, $\mathcal{H}(D)=\sum_g\|[\pi(g),\sqrt{D}]\|_2^2$, the sum over signed generators, satisfies $\mathcal{H}(D)\,(S(D)+1)^2\geq c$.

Proof. An exact N-qubit algebra with short words. Put $X_i=\pi(a^ixa^{-i})$ and $Z_i=\pi(a^iza^{-i})$, $0\leq i<N$; by Proposition 14.2 they are independent Pauli pairs, and Bob's space is $\mathbb{C}^{2^N}\otimes K_N$. For $s,t\in\{0,1\}^N$ the Pauli $P_{s,t}=X^sZ^t$ is represented by the sorted word $a^{i_1}u_{i_1}a^{i_2-i_1}u_{i_2}\cdots u_{i_k}a^{-i_k}$, $u_i=x^{s_i}z^{t_i}$, of length at most $4N$; let $M_{s,t}$ be its reversed Alice mirror, with error at most $4Ne$.

An always-valid isometry and an exact teleportation identity. Alice's map $W_A|v\rangle=d^{-1}\sum_{s,t}M_{s,t}|v\rangle\otimes|s\rangle|t\rangle$, $d=2^N$, is an isometry whatever the $M_{s,t}$ are, and by orthogonality of the labels $\|W_A\Psi-d^{-1}\sum_{s,t}P_{s,t}\Psi\otimes|s,t\rangle\|^2=d^{-2}\sum_{s,t}\|(M_{s,t}-P_{s,t})\Psi\|^2\leq16N^2e^2$. On the labels apply $K=\mathrm{CNOT}_{T\to S}(I\otimes H_T)$. For a basis input $|j\rangle$ of Bob's logical register B_0, the ideal vector $d^{-1}\sum_{s,t}(-1)^{t\cdot j}|j+s\rangle_{B_0}|s\rangle|t\rangle$ becomes $|\Phi_d\rangle_{B_0S}\otimes|j\rangle_T$, so by linearity the ideal vector is Φ_d on B_0S times the original state with B_0 relabelled T. The actual output is within $4Ne$ of it, and (2.4) gives the rates, in either direction since the output is a high-fidelity Bell state.

An exact upper bound. Let $\Gamma=P_\Omega\rtimes H_3$, $K=P_\Omega\rtimes\mathrm{FSym}(\Omega)$, $\Gamma/K=\mathbb{Z}^2$ with a,b the coordinate generators, and σ the negative spin representation of K on the incomplete tensor product of qubits stabilized by the all-zero vector: x_p,z_p act as Paulis, $J=-I$, and finite permutations permute qubits. With the section $t_q=a^{q_1}b^{q_2}$, the induced representation $\pi(g)(|q\rangle\otimes v)=|q+\lambda(g)\rangle\otimes\sigma(t_{q+\lambda(g)}^{-1}gt_q)v$ is exact, by the cocycle identity, with $\pi(J)=-I$. For clocks in the box $[-L,L]^2$ and adjacent clocks, every active cocycle has zero translation and length at most $4L+5$, so it permutes the finite set $S_L=\{C_i,A_i,B_i:0\leq i\leq 4L+8\}$ of $12L+27$ sites into itself, and the active lamps sit in S_L. Let D_0 be maximally mixed on S_L and vacuum elsewhere, $f_j=(L+1)^{-1/2}\cos(\pi j/(2(L+1)))$ for $|j|\leq L$, and $D_L=\sum_q f_{q_1}^2f_{q_2}^2|q\rangle\langle q|\otimes D_0$. Every active cocycle commutes with D_0, so $\|[\pi(a),\sqrt{D_L}]\|_2^2=\|[\pi(b),\sqrt{D_L}]\|_2^2=\sum_j(f_{j+1}-f_j)^2=2-2\cos\frac{\pi}{2(L+1)}$ and the other generators commute with $\sqrt{D_L}$. The canonical purification, with Alice's transposes, is an admissible pure input with mirror errors $O(1/L)$ and entropy $S(D_L)\leq12L+27+2\log(2L+1)$. Take $L\asymp1/e$.

The Dirichlet inequality. For a normal density D, the canonical purification with transpose mirrors has every mirror error at most $\mathcal{H}(D)^{1/2}$, and its one-way rate is $S(D)$; the extraction bound gives $S(D)+1\geq c\,\mathcal{H}(D)^{-1/2}$. □

The traces of the exact negative algebra can be described completely. Let $A=\bigotimes_{p\in\Omega}M_2$ be the negative group algebra of P_Ω, with its unique trace tr_A, and $\mathcal{B}=A\rtimes H_3$; for a finite permutation f let $V_f\in A$ be its qubit-permutation unitary and U_h the implementers in $\mathcal{B}$.

Proposition 14.4 (Negative traces)**.** *The tracial states of $\mathcal{B}$ are exactly $\tau_\chi(aU_h)=0$ for $h\notin\mathrm{FSym}(\Omega)$ and $\tau_\chi(aU_f)=\mathrm{tr}_A(aV_f)\chi(f)$ for $f\in\mathrm{FSym}(\Omega)$, where χ ranges over the normalized positive-definite class functions of $\mathrm{FSym}(\Omega)$, which Thoma classified [136]. In particular $\tau(s)=\chi(s)/2\in[-\frac12,\frac12]$ and $\tau(\gamma)=\chi(\gamma)/4\in[0,\frac14]$ for the three-cycle γ, so $\|\gamma-I\|_{2,\tau}^2\geq\frac32$, with equality for the trivial character.*

Proof. The unitaries $W_f=V_f^*U_f$ commute with A and represent FSym, so $A\rtimes\mathrm{FSym}\cong A\otimes C^*(\mathrm{FSym})$. In a tracial GNS representation the closure of A is the hyperfinite II_1 factor

R, and averaging over finite tensor factors gives $\tau(aW_f) = \mathrm{tr}_A(a)\chi(f)$ with $\chi(f) = \tau(W_f)$. An infinite-support permutation h induces an outer automorphism of R (a local Pauli at a moved point outside the support of a finite-site approximant commutes with the approximant, while its image is at L^2-distance $\sqrt{2}$), and the conditional expectation onto R intertwines, $E_R(U_h)a = \alpha_h(a)E_R(U_h)$; in a factor a nonzero intertwiner would make α_h inner, so $E_R(U_h) = 0$. Conversely, for any trace χ on $C^*(\mathrm{FSym})$ the trace $\mathrm{tr}_A \otimes \chi$ on $A \rtimes \mathrm{FSym}$ is invariant under H_3, which preserves cycle types, and its composite with the canonical expectation is tracial. The values follow from $\mathrm{tr}(\mathrm{SWAP}) = \frac{1}{2}$ and $\mathrm{tr}(V_\gamma) = \frac{1}{4}$; positivity of $\chi(T_n^2)$ for $T_n = (n-1)^{-1}\sum_{j=2}^{n}(1\ j)$ gives $\chi(\gamma) \geq 0$. □

This is an application of Thoma's classification and the standard inner-action argument, not a new classification. It shows in particular that the limiting negative traces are not unique, so no single flat native trace covers them; and since every exact finite-dimensional representation kills J, ordinary matrix stability towards exact finite models is impossible here.

14.3 The configuration source

Let R_{P} be the 27-word Pauli certificate in p, q, t, d, x, z of [101, Definition 7.1], with $J = xzxz$; the lower bound uses only these words, which are not asserted to present the configuration group G_{P}, and the upper bound uses G_{P} itself.

Theorem 14.5 (Exact configuration source)**.** *Under (EX) for R_{P}, with $\pi(J) = -I$ and $0 < e \leq \frac{1}{24}$,*

$$E_{\mathrm{D}}^{B\to A}(\rho_{AB}) \geq \tfrac{3}{4}\,2^{\lfloor 1/(24e)\rfloor}, \qquad E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \geq \tfrac{1}{2}\,2^{\lfloor 1/(24e)\rfloor}.$$

Exact induced representations of G_{P} with finite-rank pure states give an upper bound $2^{C/e}$, so both directional infima are $2^{\Theta(1/e)}$. If J is only an exact central reflection with $\|(J-I)\Psi\| \geq \Delta$, $0 < \Delta \leq 2$, then $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \geq (\Delta^2/8)\,2^{\lfloor \Delta/(48e)\rfloor}$ whenever the integer scale is at least one; this is a lower bound, not a matching joint visibility law.

Proof. Addresses of length $6N-3$. On the first ray $p(i) = i+1$. For $u = \{i_1 < \cdots < i_k\} \subseteq [N]$ put $S_u = p^{i_1-1}tp^{i_2-i_1}t\cdots tp^{1-i_k}$, of length at most $3N-2$. In the companion's notation U_i is a finite-permutation word carrying 1 to i and $\tau_i = U_i t U_i^{-1}$; the four host words of R_{P} present H_3 [101, Appendix A], $U_i^{-1}p^{i-1}$ lies in the stabilizer of the first point, which is $\langle sp, sq\rangle$ [101, Lemma 7.4], and the relators $[t, sp], [t, sq]$ of R_{P} make t commute with it. So $p^{i-1}tp^{1-i} = \tau_i$ in every exact certificate representation, and multiplying in increasing i gives $S_u = T_u$, the companion's transporter [101, Definition 7.5]. By [101, Theorem 7.8 and Lemma 7.11], read exactly, $X_u = \pi(S_u x S_u^{-1})$ and $Z_u = \pi(S_u z S_u^{-1})$ are $M = 2^N$ independent Pauli pairs, with individual words of length at most $6N-3$.

Lower bounds. Theorem 14.1 with $r = 6N-3$ and $N = \lfloor 1/(24e)\rfloor$: the argument $p_N = ((6N-3)e)^2/4 \leq \frac{1}{64}$ gives the reverse bound with $1 - 2h(\frac{1}{64}) > \frac{3}{4}$, and $((6N-3)e)^2/2 \leq \frac{1}{32}$ gives the forward bound with $1 - 2h(\frac{1}{32}) > \frac{1}{2}$. For the visibility version, apply the full lift of Theorem 6.10; Alice's transposed central projection has the same weight $p_A = \mathrm{Tr}(RP_-) \geq \Delta^2/4$ as Bob's, she measures it and announces success, conditional mirror errors are at most $\sqrt{2}e/\sqrt{p_A}$, and the per-bit argument with $N = \lfloor\sqrt{p_A}/(24e)\rfloor$ gives $\frac{p_A}{2}2^N \geq \frac{\Delta^2}{8}2^{\lfloor\Delta/(48e)\rfloor}$. This is an equality of marginal sector weights, not of sector outcomes.

Upper bound. By [101, Proposition 6.8 and (5)], $G_{\mathrm{P}} = E_V \rtimes A$ has $G_{\mathrm{P}}/K = \mathbb{Z}^2$ with $K = E_V \rtimes (V \rtimes \mathrm{GL}_{\mathrm{fin}}(Y))$, $\lambda(p) = (1,0)$, $\lambda(q) = (0,1)$. Induce the negative spin representation of K on the incomplete tensor product of qubits indexed by all configurations, with the section $p^{c_1}q^{c_2}$; it is exact with $\pi(J) = -I$. For clocks in $[-L, L]^2$ the clock lemma of the companion [101, Lemma 9.3], with $m = L+1$, $r = R = 2L+5$ and trivial ray swap,

puts every active cocycle in the subgroup preserving $W_R = \mathbb{F}_2^{(Y_R)}$, the configurations of the $3R$ points of height at most R. With D_0 maximally mixed on the $|W_R| = 2^{6L+15}$ qubits of W_R and the cosine clock as before, the mirror errors are at most $\pi/(2(L+1))$ and $S(D_L) \le 2^{6L+15} + 2\log(2L+1)$. Take $L = \lceil \pi/(2e) \rceil$. □

The minimum rank of such a density is doubly exponential: its double logarithm is $\Theta(1/e)$.

14.4 Subset lamps: every positive integer exponent

Fix $k \ge 1$. Let Y be the three rays of H_3, with p, q acting as in [101, Definition 6.1], let $X_k = \binom{Y}{k}$ be the set of unordered k-subsets, and let E_{X_k} be the central Pauli group with finite-support normal forms (ϵ, a, b) over $\mathbb{F}_2$ and product $(\epsilon, a, b)(\epsilon', a', b') = (\epsilon + \epsilon' + \langle b, a' \rangle, a + a', b + b')$, with common centre $J = (1, 0, 0)$. Put $G_k = E_{X_k} \rtimes H_3$.

Proposition 14.6 (Presentation)**.** *G_k is presented on p, q, x, z by $6k + 11$ relators: the four relators of Johnson's presentation of H_3 in p, q [77], [84, Theorem 2.8] (in the form [101, (6)]), with $s = [p, q]$); x^2, z^2, $(xz)^4$; the commutators of x and z with $s_i = p^{i-1} s p^{1-i}$ $(1 \le i < k)$, $v_k = p^k q p^{-k}$ and $u_k = v_k q^{-1} p$, which generate the setwise stabilizer of $[k]$; the $4k$ commutators $[a, h_r b h_r^{-1}]$, $a, b \in \{x, z\}$, $h_r = p^{k-r}$, $0 \le r < k$; and $[J, p], [J, q]$ with $J = (xz)^2$. For $k = 2$ the literal list has 23 freely reduced words, of maximum length 22 and total length 252.*

Proof. The words s_i act as adjacent transpositions of $[k]$, and u_k, v_k fix $[k]$ pointwise and act on its three-ray complement as the standard generators of H_3; so they generate the setwise stabilizer $S_k \times H_3$. Ordered pairs of distinct k-subsets are classified up to H_3 by the size r of their intersection, because H_3 contains all finitary permutations, and $([k], p^{k-r}[k])$ represents size r. Without the two central relators the presented group $\hat{G}$ maps onto $W = (\bigoplus_{X_k} D_4) \rtimes H_3$; conversely the stabilizer relators make the transported pairs well defined, the prototypes make distinct sites commute, and this gives a homomorphism $W \to \hat{G}$ inverse to the first on generators. Adding $[J, p], [J, q]$ identifies all onsite centres, so the normal closure is the even-centre subgroup, and the quotient is G_k; its normal form shows $J \ne 1$. The input from the literature is Johnson's presentation; this is the permutational-wreath mechanism of Cornulier [48], written out. □

Theorem 14.7 (Integer exact-source spectrum)**.** *Under (EX) for G_k, both directional infima are $\Theta_k(e^{-k})$. Explicitly, with $r_N = 2D_k(N-k) + 1$, $D_1 = 1$ and $D_k = 6k - 7$ for $k \ge 2$, both rates are at least $\binom{N}{k}[1 - 2h(r_N^2 e^2/2)]$ when the argument is at most $\frac{1}{2}$ (the reverse estimate may use $r_N^2 e^2/4$); for $k = 2$ and $e \le \frac{1}{160}$ this is at least $e^{-2}/51200$. The upper bound uses all k-subsets of a common three-ray point region of height $R_L = 4L + \max(k, 2)$, flat there and vacuum outside, and a smooth two-dimensional clock, with entropy at most $\binom{3R_L}{k} + 2\log(2L+1)$ and root generator error at most $\pi/[2(L+1)]$.*

Proof. Addresses. Let $g_0 = p$ and, for $m \ge 1$, $g_m = u_m$, which fixes $[m]$ and moves the rest of the first ray up by one. For $S = \{a_1 < \cdots < a_k\} \subseteq [N]$, $a_0 = 0$, the word $T_S = \prod_{i=1}^{k} g_{i-1}^{a_i - a_{i-1} - 1}$, factors in increasing i, carries $[k]$ to S and has length at most $D_k(N-k)$; so each of the $\binom{N}{k}$ transported pairs has individual words of length at most r_N. On an exact negative representation these sites generate $M_{2^{\binom{N}{k}}}$. Theorem 14.1 gives the lower bounds; for $k = 2$, $N = \lfloor 1/(40e) \rfloor$ gives $r_N e \le \frac{1}{4}$ and the stated constant.

Upper bound. The kernel of $G_k \to H_3 \to \mathbb{Z}^2$ is $E_{X_k} \rtimes \mathrm{FSym}(Y)$; represent it on the finite-excitation spin space of the subset sites, with point permutations permuting subset sites (they may move infinitely many sites but fix the all-zero vector), and induce with the section $p^{c_1} q^{c_2}$. For a signed p or q edge with an endpoint in $[-L, L]^2$ the cocycle has point-word length at most $4L + 2$ and fixes every point above that height, and the transported base lamp

at an active clock has height at most $2L + k$; so one common region of height R_L contains all supports. A finitary point permutation may move infinitely many subset sites outside $\binom{S_L}{k}$ but preserves this set and its complement, so the flat density on its $\binom{3R_L}{k}$ qubits with vacuum outside commutes with every active cocycle. With the cosine clock, the p, q errors are $\sqrt{2 - 2\cos(\pi/(2(L+1)))}$ and the x, z errors vanish; pure concentration attains the entropy in either direction. □

Theorem 14.8 (Joint visibility law)**.** *On the exact relations of G_k, let J be a central reflection, not necessarily $-I$ on the whole input, with $\|(J-I)\Psi\| \geq \Delta$, $0 < \Delta \leq 2$, and actual generator mirrors at error e. Uniformly for $e \leq c_k\Delta$,*

$$\inf E_{\mathrm{D}}^{A\to B} = \Theta_k(\Delta^{k+2}e^{-k}), \qquad \inf E_{\mathrm{D}}^{B\to A} = \Theta_k(\Delta^{k+2}e^{-k}).$$

Proof. The negative weight $p = \operatorname{Tr}\rho_B \frac{1}{2}(I - J)$ is at least $\Delta^2/4$. For $B \to A$ Bob filters the negative sector and announces; for $A \to B$ Alice applies the full lift of Theorem 6.10 and filters her exact lifted central sector, whose marginal weight is also p. The conditional mirror errors are at most $e/\sqrt{p}$, or $\sqrt{2}e/\sqrt{p}$ in the forward case, and $p\binom{N}{k}[1 - 2h(r_N^2e^2/(2p))]$ with $N \propto \sqrt{p}/e$ gives $c\Delta^{k+2}e^{-k}$. For the upper bound take the locally flagged mixed input that, with probability $p = \Delta^2/4$, is a pure negative-sector witness of Theorem 14.7 at error $e/\sqrt{p}$, and otherwise a trivial positive-sector product state; the flag is available to both parties. Squared mirror errors are multiplied by p, the visibility is exactly $2\sqrt{p} = \Delta$, and both directional rates equal p times the negative entropy, with no $h(p)$ term, by local flag reading and pure concentration. □

The lower bound uses "weight at least", not an equality; Alice's and Bob's flag outcomes are not assumed identical on an arbitrary input; and the law concerns mixed-state E_{D}, not a matching law for the marginal entropy or rank, which carry $h(p)$ and no factor p, respectively.

14.5 A compact exponential source by replication

Theorem 14.9 (A five-generator exponential source)**.** *Let G_1 be the case $k = 1$ of Proposition 14.6, the exact point source on p, q, x, z with 17 relators. The map $\varphi(p) = p^3$, $\varphi(q) = q^3$, $\varphi(x) = \prod_{i=0}^{2} p^i x p^{-i}$, $\varphi(z) = \prod_{i=0}^{2} p^i z p^{-i}$, $\varphi(J) = J$ is an injective endomorphism, and its ascending HNN extension $\Gamma_3 = \langle G_1, t \mid tgt^{-1} = \varphi(g)\rangle$ has five generators and 21 freely reduced relators, of maximum length 22 and total length 200. Radix words address N independent sites in individual length at most $8\lceil \log_3 N\rceil + 1$, and under (EX), for $e \leq \frac{1}{64}$, both directional rates are at least $\frac{1}{2}\,3^{\lfloor 1/(64e)\rfloor}$. Exact induced representations over $\mathbb{Z}[1/3]^2 \rtimes \mathbb{Z}$ with finite-rank pure states give both directional infima $\exp(\Theta(1/e))$.*

Proof. Replication. Identify $Y \times [3]$ with Y by $I(1, j, r) = (1, 3(j-1) + r)$ and $I(i, j, r) = (i, 3(j-1) + 4 - r)$ for $i = 2, 3$; the reversed copy order on the donor rays is essential. Then $\varphi_H(g) : I(y, r) \mapsto I(gy, r)$ is an injective endomorphism of H_3 with $\varphi_H(p) = p^3$, $\varphi_H(q) = q^3$, including the crossing endpoints. On lamps φ replicates each site on its block; the onsite commutator becomes $J^3 = J$ because 3 is odd, and the normal forms show injectivity. By the normal-form theorem for ascending HNN extensions G_1 embeds in Γ_3, and J stays central and nontrivial. An even replication would kill the onsite commutator.

Addresses and the lower bound. With W_0 empty and $W_n = p^r t W_a t^{-1}$ for $n = 3a + r$, $0 \leq r \leq 2$, W_n represents p^n and $|W_n| \leq 4\lceil \log_3(n+1)\rceil$; the first N point lamps have words $W_{j-1}xW_{j-1}^{-1}$, $W_{j-1}zW_{j-1}^{-1}$ of length at most $r_N = 8\lceil \log_3 N\rceil + 1$. With $d = \lfloor 1/(64e)\rfloor$ and $N = 3^d$, $r_N e \leq \frac{9}{64}$, and Theorem 14.1 gives at least $\frac{1}{2}N$ in both directions.

Upper bound. The kernel K of $\Gamma_3 \to \mathbb{Z}[1/3]^2 \rtimes \mathbb{Z}$ is the direct limit of $K_0 = E_Y \rtimes \mathrm{FSym}(Y)$ under φ, locally finite. At stage L put $p_L = t^{-L}pt^L$, $q_L = t^{-L}qt^L$, and for R let $K_{L,R} \cong$

$E_{3R} \rtimes \mathrm{Sym}_{3R}$, of order $2^{6R+1}(3R)!$, be the subgroup generated by the lamps and finitary permutations of the three rays up to height R. Take $M = (L+1)3^{2L}$, $R = 8M + 8$, the section $s(a, b, j) = t^j p_L^a q_L^b$ on $|a|, |b| \le M$, $|j| \le L$, extended consistently to the neighbours. Every edge with an endpoint in the support has its cocycle in $K_{L,R}$: $t^{\pm 1}$ and $p^{\pm 1}$ edges have trivial cocycle, $q^{\pm 1}$ edges a finitary word fixing points above height $8M$, and x, z edges blocks of original lamps of height at most $3M$; inward and outward boundary edges are included. Induce the negative-J part of the regular representation of $K_{L,R}$, with D_0 uniform on it, and put cosine amplitudes $f_M(a) f_M(b) f_L(j)$ on the clocks. The t error is at most $\pi/[2(L+1)]$, the p, q errors at most $3^{2L}\pi/[2(M+1)] \le \pi/[2(L+1)]$ by telescoping shifts, and the entropy at most $6R + \log((3R)!) + 2\log(2M+1) + \log(2L+1) \le e^{CL}$. Take $L = \lceil \pi/(2e) \rceil$. □

Exponential order was already obtained by Theorem 14.5; this is a smaller alternative mechanism, compressing addresses by replication instead of configurations, not a new growth exponent. Exact address length and approximate relation filling are different resources: the scaling relator alone, $tpt^{-1} = p^3$ with $t = 1$ and $p = e^{i\pi/(3^k-1)}$, has base defect $O(3^{-k})$ and iterated defect $|t^k p t^{-k} - p^{3^k}| = 2$. The inputs to Proposition 14.6 and Theorems 14.7 and 14.9 from the literature are Johnson's presentation of H_3, Cornulier's criterion and the standard HNN and induction machinery; the operational inputs are Theorems 6.10 and 14.1.

IV. Limits

15 Obstructions and refutations

This section records what does *not* hold, or what a proof cannot do, with proofs. Several entries refute statements we once expected; they are kept because they mark the boundary of the method.

15.1 Positivity under partial transposition is not distillability

Theorem 15.1 (A PPT gap)**.** *Let a finite binary linear system with nonempty rows be classically inconsistent, let every incidence be tested with probability at least $p_{\min} > 0$, and choose a parity-contradiction certificate, a set of rows in which every variable occurs an even number of times and whose signs multiply to -1, with total number N of variable occurrences. Every finite-dimensional strategy on a state with positive partial transpose has loss at least $4p_{\min}/(N^3 + 3N^2 + N)^4$.*

Proof. Relabelling illegal answers and local Naimark dilations with pure ancillas preserve PPT and do not increase the loss. Let A_{ij} be Alice's commuting row reflections, $\prod_{j \in V_i} A_{ij} = b_i I$, and B_j Bob's, so $\|(A_{ij} - B_j)\sqrt{\rho}\|_2 \le d = 2\sqrt{\varepsilon/p_{\min}}$, with ε the loss and the norm taken with the joint density. Choose one row reflection A_j per variable. If $\sigma = \rho^{T_B} \ge 0$, then B_j^T mirrors A_j in σ with the same error, since $\operatorname{Tr} \sigma(A_j \otimes I - I \otimes B_j^T)^2 = 2 - 2\operatorname{Tr} \rho(A_j \otimes B_j)$. For $W = A_j A_k \otimes B_j B_k$, $|\operatorname{Tr} \rho W - 1| \le 2d$; $\operatorname{Tr} \rho W = \operatorname{Tr} \sigma(A_j A_k \otimes B_k^T B_j^T)$, and transferring the two Bob factors to Alice in σ costs $2d$ and gives $A_j A_k A_j A_k$, whose expectation is the same in ρ and σ. So $\Re \operatorname{Tr} \rho A_j A_k A_j A_k \ge 1 - 4d$ and $\|[A_j, A_k]\sqrt{\rho}\|_2^2 \le 8d$. For a certificate row i with representative word R_i, replacing its letters by the row reflections costs at most $r_i(r_i+1)d$, so $\|(R_i - b_i)\sqrt{\rho}\|_2 \le r_i(r_i+1)d$; telescoping against scalars, the ordered product P of the certificate words has $\|(P+I)\sqrt{\rho}\|_2 \le (N^2+N)d$. Every letter of P occurs an even number of times; sorting with at most $N^2/2$ adjacent interchanges, each costing at most $\sqrt{8d} + 2Nd$ with its suffix moved to Bob, gives $\|(P-I)\sqrt{\rho}\|_2 \le \sqrt{2}N^2\sqrt{d} + N^3 d$. Hence

$2 \le (N^3+N^2+N)d + \sqrt{2}N^2\sqrt{d}$, which for $d \le 1$ is at most $(N^3+3N^2+N)\sqrt{d}$, and the bound follows; if $d > 1$ the loss exceeds $p_{\min}/4$. □

Sufficiently good near-winners of such a game are therefore entangled with negative partial transpose. This alone proves neither distillability nor any growing rate, and it does not bypass the NPT distillability problem; PPT relaxations of device-independent entanglement constraints are due to Moroder et al. [103]. The operational theorems of Sections 6 and 7 are separate proofs.

15.2 No exact output from full-rank mixed states

Remark 15.2 (No exact entangled output; standard). If ρ_{AB} is a full-rank finite-dimensional state, no nonzero stochastic separable operation on finitely many copies, and hence no branch of an LOCC protocol, outputs an exact entangled pure state with nonzero probability. Every near-winning strategy can be made full rank by an arbitrarily small depolarization at an arbitrarily small increase of loss.

Proof. If $\rho \ge \lambda I$ with $\lambda > 0$ and $\Lambda(\rho) = p|\Phi\rangle\langle\Phi|$ for a separable completely positive Λ, then $0 \le \lambda\Lambda(I) \le p|\Phi\rangle\langle\Phi|$; a positive separable operator supported on an entangled line is zero, so $\Lambda(I) = 0$ and $\Lambda = 0$. The same applies to $\rho^{\otimes n}$. For the second claim mix a strategy of loss $\varepsilon/2$ with weight $t \le \varepsilon/2$ of the maximally mixed state. □

So the exact-output statements for pure strategies (Theorem 11.7) and exactly synchronous strategies (Theorem 6.11) cannot be extended to arbitrary mixed near-winners; approximate extraction is the theorem. This is standard finite-copy logic, not a new no-distillation result.

15.3 Restricted universal linear compilers

Proposition 15.3 (Two exclusions)**.** (a) *There is no split finite linear envelope ($\sigma F = I_N \otimes \mathrm{id}$, with σ into M_N of the negative quotient) in which every source generator is recovered as a fixed scalar times* one *group word in the involutive constraint generators.* (b) *A finite group K with central involution J whose negative group algebra is a single matrix algebra M_d has negative character values in $\{0, \pm d\}$, so it cannot contain a coordinate transposition of four sectors, of trace two, as a group-word anchor.*

Proof. (a) Take the source $\mathbb{Z} \times C_2$ with the negative central involution and its characters $u \mapsto z$, $z \in \mathbb{T}$. Composing σ with a character gives an N-dimensional representation of the game algebra in which every involutive generator has determinant ± 1, so a fixed scalar λ times a fixed word has determinant $\pm\lambda^N$; but $\sigma F(u) = I_N \otimes u$ has determinant z^N. So $z^N \in \{\pm\lambda^N\}$ for every z, which is impossible. (b) The tensor cube of the unique negative irreducible representation is again negative, so $\chi^3 = d^2\chi$ pointwise. □

A tested polynomial *minority-frame* decoder, which recovers each source generator in the corner selecting the minority coordinate of four commuting involutions with product -1, also fails: tensoring three lifted models of the source $\langle u, v, q, w \mid uvq^{-1}, w^2, [w, \cdot]\rangle$ gives an exact 128-dimensional model of every frame relation in which the decoded relator uvq^{-1} equals $-I_{32}$ on one sector. This refutes that candidate at zero model error, not arbitrary polynomial decoders. Theorem 5.6 produces Boolean constraint games, and the restricted grammar of Theorem 11.3 survives; neither establishes nor rules out an unrestricted linear compiler (Section 18).

15.4 Raw pair bounds below the exponent three halves

Proposition 15.4 (Raw pairs). *For every $M \geq \lceil 384\pi \rceil$ there is a native model $\mathcal{N}_M$ of Section 13.5, namely the model of [55, Proposition 6.5] with the spanning trees chosen in the proof, such that on its flat maximally entangled purification, with all actual word mirrors exact and source squared error at most $14.26M^{-3}$, the raw lamp commutator satisfies $\|[x, a^{4M} z a^{-4M}] - I\|_2^2 \geq 1/(128\pi)$ in the normalized Hilbert–Schmidt norm. Consequently no estimate $\|([x, a^n z a^{-n}] - I)\Psi\| \leq C n^u e$ with $u < \frac{3}{2}$ holds uniformly over models and distances; flag dilution, weight p on such a block and $1-p$ on a more accurate one, keeps this failure at an arbitrarily small critical budget $N^3 e^2 \leq C' p$, with ratio to $N^{2u} e^2$ growing like M^{3-2u}.*

Proof. The model. We use the data of [55, Section 6] in its conventions: $\ell = 2M + 2$, the reduced colour coordinates $1, \dots, L$ with $L = 4M + 4$, clocks $(i, j) \in (\mathbb{Z}/M\mathbb{Z})^2$, the reverse cycles $c_r = (1\ r\ r{-}1\ \cdots\ 2)$, the base a-links $A_{i,j} = c_{\ell-i-j}$, the base b-links, and the two exact moves η_1, on the a-link at $(M-1, M-1)$, and η_2, on a b-link. The plaquette (i, j) has bottom and top edges the a-links at (i, j) and $(i, j+1)$, and left and right edges the b-links at (i, j) and $(i+1, j)$. The regions are those of [55, proof of Proposition 6.5]: for $5 \leq t \leq M + 3$ the bulk region R_t, with source $(M-1, M-1)$, consists of its source, the column plaquettes $(M-1, j)$ with $1 \leq j \leq M-2$, and the staircase plaquettes (i, j) with $0 \leq i \leq M-3$, $0 \leq j \leq M-2$, $i + j \leq t + M - 8$; the region T, with source $(0, M-2)$, consists of its source and the plaquettes with $i + j \leq M - 3$; signs and labels are as given there. These regions satisfy the four regional hypotheses of [55, Lemma 6.4], every label at the initial vertex of an interior edge has both coordinates at least 5 [55, proof of Proposition 6.5], and $|R_t| \geq M(M-1)/2$, $|T| = 1 + (M-1)(M-2)/2$. Lemma 6.4 of [55] solves $\operatorname{curl} \varphi_R = F_R - s_R \pi \delta_{q_0}$, for the constant fields $F_{R,q} = s_R \pi / |R|$, along an *arbitrary* spanning tree of the plaquettes of R joined through interior edges, rooted at the source, with zero potential off the tree. The resulting H4 holonomies depend only on the fields, so the bound $\|\mathrm{H4} - I\|_2^2 \leq 14.26 M^{-3}$ of [55, Proposition 6.5] holds for every tree; and the rotations act in the label planes, which avoid the coordinates $1, \dots, 4$, so the other Houghton relators [55, Lemma 6.4] and the lamp relators (Section 13.5) stay exact whatever tree is used. On the flat maximally entangled purification with Alice's transposes every mirror is exact and every state commutator vanishes, so the model has precision $e_M \leq \sqrt{14.26}\, M^{-3/2}$.

Row-spine trees. For R_t: on each row $1 \leq j \leq M-2$ join the staircase plaquettes $(0, j), (1, j), \dots$ consecutively through their common b-links, and join $(0, j)$ to the column plaquette $(M-1, j)$ through the b-link at $(0, j)$, across the seam; join the column plaquettes consecutively through the a-links at $(M-1, j)$, $2 \leq j \leq M-2$, and $(M-1, M-2)$ to the source through its bottom edge. Row 0 cannot cross the seam, since $(M-1, 0) \notin R_t$: join its plaquettes to $(0, 0)$ through b-links, and $(0, 0)$ to $(0, 1)$ through the a-link at $(0, 1)$. For T: join each row to its plaquette $(0, j)$ through b-links, and the plaquettes $(0, 0), \dots, (0, M-3)$ and the source $(0, M-2)$ through the a-links at $(0, j)$, $1 \leq j \leq M-2$. These edges are interior, the trees are spanning and acyclic, and each source is a leaf attached by its bottom edge, as the third regional hypothesis of [55, Lemma 6.4] requires.

Angles. If a tree edge e separates a subtree S not containing the source from the rest, summing the curl equation over S, in which every edge inside S cancels and every other edge but e carries zero potential, gives $|\varphi_R(e)| = \pi |S| / |R|$. Below the spine edge at $(M-1, j)$, $j \geq 2$, lie the column plaquettes $(M-1, 1), \dots, (M-1, j-1)$ and the staircase rows $0, \dots, j-1$, at most $j(M-1)$ plaquettes, so $|\varphi_{R_t}| \leq 2\pi j / M$; below the T-spine edge at $(0, j)$ lie at most $j(M-2)$ plaquettes, so $|\varphi_T| \leq 2\pi j / (M-1)$; the edge at $(0, 1)$ of R_t carries an angle at most $2\pi / M$. On a row $j \geq 2$ the only a-links with nonzero potentials are thus those at $(M-1, j)$, carrying the rotations of the bulk regions, and at $(0, j)$, carrying that of T. The bulk rotations act in disjoint planes, so the first link factor is within $\max_t |\varphi_{R_t}| \leq 2\pi j / M$ of I in operator

norm, not within the sum, and the second within $2\pi j/(M-1)$. Neither η-move lies on an a-link of a row $2 \le j \le M-2$.

The unperturbed circuit. In the representation of [55, Section 6.2] the generator a is the adjoint of the right-action operator, so a moves the clock by $-e_1$ and applies the inverse of the link at the new clock to the one-particle space $\mathbb{C}^L$. From the clock $(0,j)$ the M-step circuit along row j therefore applies, in chronological order, the forward cycles $d_r = c_r^{-1} = (1\ 2\ \cdots\ r)$ of lengths $r = d+1, d+2, \dots, R$, where $d = M+2-j$ and $R = M+d$. Following the points,

$$P(v) = \begin{cases} v+M, & 1 \le v \le d, \\ R-v+1, & d < v \le R, \\ v, & v > R : \end{cases}$$

for $v \le d$ every cycle increments v; for $d < v \le R$ the cycles shorter than v fix it, the cycle of length v sends it to 1, and the remaining $R-v$ cycles increment it. For $j \ge 3$, $d < M$, so P exchanges the low block $[1,d]$ with the high block $[M+1,R]$, P^2 reflects each of them, and P reflects the middle block $[d+1,M]$; hence $P^4 = I$. A circuit from another clock of the row applies a cyclic rotation of the same factors, which is conjugate to P and has fourth power I. The orientation matters: the same cycles applied in the opposite chronological order, or their inverses in the same order, do not in general give an element of order four. The order used here is that of the single unitary a; its inverse gives the inverse circuits, with the same property.

The perturbed circuit. The one-particle holonomy $W_{i,j}(a^{4M})$ of a^{4M} from a clock (i,j) of row j uses every link of the row four times, and the rotated links differ from the base links only in the columns $M-1$ and 0. Telescoping against the unperturbed product, which is I, gives

$$\|W_{i,j}(a^{4M}) - I\|_{\mathrm{op}} \le 4\Big[\frac{2\pi j}{M} + \frac{2\pi j}{M-1}\Big] \le \frac{24\pi j}{M},$$

uniformly in i. For $3 \le j \le \lfloor M/(64\pi) \rfloor$ this is at most $\frac{3}{8}$, so $|W_{i,j}(a^{4M})_{1,1}| \ge \frac{1}{2}$; for $M \ge \lceil 384\pi \rceil$ these rows number at least $M/(128\pi)$, so they carry a fraction at least $1/(128\pi)$ of the clocks. This is a one-particle estimate; it is not claimed that fermionic quantization keeps operator norms small on the whole space.

A Gaussian identity. The lamps are $x = i\gamma_{1,1}\gamma_{1,2}$ and $z = i\gamma_{1,2}\gamma_{1,3}$, where $\gamma_{1,1}, \gamma_{1,2}$ are the Majorana operators of the first flavour at the coordinate 1 and $\gamma_{1,3}$ one of the second flavour there. At each clock $a^n z a^{-n}$ is the conjugate of z by the fermionic quantization of the one-particle holonomy W of a^n there, or of its inverse, which has the same diagonal moduli. This conjugation sends $\gamma_{1,2}$ to a unit real combination A of first-flavour Majorana operators and $\gamma_{1,3}$ to a unit real combination B of second-flavour ones. Write $A = A_\parallel + A_\perp$ with $A_\parallel$ in the span of $\gamma_{1,1}, \gamma_{1,2}$; its squared coefficient norm is $|W_{1,1}|^2$, so $A_\parallel^2 = |W_{1,1}|^2 I$. Conjugation by x reverses $A_\parallel$ and fixes $A_\perp$ and B, so $x(iAB) - (iAB)x = -2iA_\parallel Bx$ and, since $B^2 = I$,

$$\|x(iAB) - (iAB)x\|_2^2 = 4|W_{1,1}|^2.$$

The normalized Hilbert–Schmidt norm is unitarily invariant, so averaging over clocks gives the exact identity $\|[x, a^n z a^{-n}] - I\|_2^2 = 4\,\mathbb{E}_c |W_c(a^n)_{1,1}|^2$, and at $n = 4M$ the good rows give at least $4 \cdot \frac{1}{4} \cdot \frac{1}{128\pi}$.

Consequences. With $n = 4M$, $n^u e_M \le 4^u \sqrt{14.26}\, M^{u-3/2} \to 0$ for $u < \frac{3}{2}$, while $\|([x, a^n z a^{-n}] - I)\Psi\|$, the normalized norm on the flat purification, stays at least $(128\pi)^{-1/2}$. For the dilution take the direct sum, with weights p and $1-p$, of this model and a native model $\mathcal{N}_{M'}$ with $(1-p)e_{M'}^2 \le p e_M^2$, each on its maximally entangled purification: the operators are block diagonal, the mirrors stay exact, the state commutators vanish, $e^2 \le 2pe_M^2 \le 29pM^{-3}$,

and the squared raw residual is at least $p/(128\pi)$. With $N = 4M$, $N^3e^2 \le 64 \cdot 29\,p$, while the squared raw residual divided by $N^{2u}e^2$ is at least $M^{3-2u}/(128\pi \cdot 29 \cdot 16^u)$. □

This refutes a uniform raw estimate, not the endpoint $u = \frac{3}{2}$, a decoder adapted to another architecture, or a suitably weighted aggregate theorem; it is not legitimate to call the endpoint settled. The proposition is stated for the row-spine trees; we do not need, and do not claim here, that the return amplitudes are the same for every legal tree.

15.5 The finite-table gate and what refutes its cheap forms

Proposition 15.5 (A finite-table gate)**.** *Let $G = \mathbb{F}_2^N \times \mathbb{F}_2^N$, $\omega(p,q) = (-1)^{t_p \cdot s_q}$ for $p = (s_p, t_p)$, let U_p be arbitrary Bob unitaries and M_p arbitrary Alice unitaries, and put*

$$\mu^2 = \mathbb{E}_q \|(M_q - U_q)\Psi\|^2, \qquad \Delta_N = \mathbb{E}_{p,q} \|(U_p U_q - \omega(p,q) U_{p+q})\Psi\|^2.$$

There are product local isometries with Bell overlap at least $1 - 2\mu^2 - 2\Delta_N$, and vector distance at most $\sqrt{2\mu^2 + 2\Delta_N}$ from a Bell state of dimension 2^N times a residual purification. If $p_0 = 2\mu^2 + 2\Delta_N < \frac{1}{2}$, then $E_{\mathrm{D}}^{A\to B} \ge (1 - 2p_0)N - h(p_0)$.

Proof. Bob's isometry $V_B|v\rangle = 2^{-N} \sum_p U_p|v\rangle \otimes |p\rangle$ is always valid. On the fresh label register $R_q|r\rangle = \omega(r+q, q)|r+q\rangle$ is an exact projective regular representation, $R_q R_r = \omega(q,r) R_{q+r}$, so it splits as N standard qubits times multiplicity. The coefficient of $R_q V_B|v\rangle$ at label p is $\omega(p,q)U_{p+q}|v\rangle/2^N$, so $\|(V_B U_q - R_q V_B)\Psi\|^2 = \mathbb{E}_p \|(U_p U_q - \omega(p,q)U_{p+q})\Psi\|^2$, and $\mathbb{E}_q \|(M_q - R_q)V_B\Psi\|^2 \le 2\mu^2 + 2\Delta_N$. The coherent Alice-label construction of Theorem 14.3, applied to the exact R_q and the individual M_q, uses only this average. □

For the raw sorted point-source table, words of length at most $4N$ give $\mu^2 \le 16N^2e^2$, so the gate reduces to controlling the same-state Δ_N; it does not assume that perfect mirrors imply Pauli multiplicativity. In an exact negative representation $\Delta_N = 0$ and Theorem 14.3 is recovered.

Proposition 15.6 (What the collective witness refutes)**.** *On the collective models of Proposition 13.9, with $e_M = \Theta(M^{-2})$, pure cost $(4 + o(1))M$ and all word mirrors exact, the sorted table satisfies $\Delta_N \ge 1 - D_M/2^N$. Consequently no bound $\Delta_N \le CN^\alpha e^2$ with $\alpha < 4$, and no modulus tending to zero with N^3e^2, holds on all actual finite point-source tuples; no Bell-error-density law $b_N \le \omega(N^3e^2)$ with $\omega(t) \to 0$ and the hashing guarantee holds; and the raw aggregate estimate $E_B \le CN^4e^2$, inserted in Theorem 6.6, is false. The quartic* table *estimate $\Delta_N \le CN^4e^2$ is not refuted.*

Proof. With exact mirrors the cross term vanishes and the Bell failure of Proposition 15.5 is at most Δ_N, while a local channel applied to a pure state of Schmidt rank D_M has overlap at most $D_M/2^N$ with Φ_{2^N}. At $N = 16M$ the lower bound tends to one, while $N^\alpha e_M^2 = O(M^{\alpha-4}) \to 0$. A density law would give a rate $N(1 - o(1))$ at $N = 16M$, exceeding the original $(4 + o(1))M$ ebits; the raw aggregate estimate would give $b_N = O(1/N)$ since $M_2 = 0$, with the same contradiction. □

These refutations do not exclude the matching quarter lower bound for the game.

15.6 Corrections that keep the resource must be accurate

Proposition 15.7 (Correction costs)**.** *Apply a Bob-local isometry, or an explicitly reversible local lift, to the original collective state of Proposition 13.9, keeping Alice's actual mirrors, and suppose the target signed generators compare on that state within weighted error η. If the target is an exact negative representation, then $\eta = \Omega(M^{-1}) = \Omega(\sqrt{e_M})$, with infinite*

target dimension allowed. If the target lies in the enlarged native class of Theorem 13.12, then $\eta = \Omega(M^{-3/2}) = \Omega(e_M^{3/4})$, conditional on that theorem. For the finite soft configuration models, of pure cost $O(e^{-2}\log(1/e))$ with exact mirrors, a correction to an exact negative representation needs $\eta = \Omega(1/\log(1/e))$.

Proof. The lifted state has the same directional rates, and since the old mirrors are exact, fixed word comparisons cost at most their length times η: the lift is a model of the target at precision $O(e_M+\eta)$ with mirror errors η. For an exact target, Theorem 14.3 forces $S_M+1 \geq c/(e_M+\eta)$ with $S_M = \Theta(M)$; for the native target Theorem 13.12 forces $S_M+1 \geq c(e_M+\eta)^{-2/3}$. For the configuration models Theorem 14.5 forces $S_e \geq \frac{3}{4}2^{\lfloor 1/(24\eta)\rfloor}$, so $\eta \geq 1/[24(1+\log(4S_e/3))]$. □

These obstruct only the stated correction contract, not replacing the state, buying more entanglement, or qualitative comparisons with vanishing error.

15.7 Restricted phase schemes

Proposition 15.8 (Scalar Weyl phases)**.** *Suppose actual unitaries on an original purification satisfy $\|(U_A-U_B)\Psi\| \leq \delta$, $\|(V_A-V_B)\Psi\| \leq \delta$ and $\|(U_A^*V_AU_A-\lambda V_A)\Psi\| \leq \delta$ for a* fixed scalar *$\lambda \in \mathbb{T}$, and a game supplies $\delta \leq C_G\sqrt{\varepsilon}$. For a reduced p/q with $\zeta = e^{2\pi ip/q}$, $\Delta = |\lambda-\zeta|$, $2 \leq N \leq q$, $m = \lfloor q/N \rfloor$, $P = mN/q$ and $\eta = \sqrt{2q\delta} + N^{3/2}\delta + \sqrt{3qN\delta + qN\Delta}$, there is a one-way protocol succeeding with probability exactly P whose output has singlet fraction at least $1-\eta^2/P$. Consequently: for every irrational phase, $E_{\mathrm{D}}^{A\to B} \geq c\log(1/\varepsilon_n)$ along a sequence $\varepsilon_n \to 0$; under polynomial growth of the continued-fraction denominators, for every small ε; and so for every algebraic λ that is not a root of unity. A finite commuting-scalar rational semialgebraic gadget whose admitted phases avoid all roots of unity admits only finitely many phases, all algebraic.*

Proof. All operators are finite-dimensional, opposite-party operators commute, and commuting unitaries of the two parties have a positive joint spectral measure on Ψ. Angles are in turns.

Random cuts across the parties. For t uniform in $[0,1/q)$ let $I_k(t) = [t+k/q, t+(k+1)/q)$, $k \in \mathbb{Z}/q\mathbb{Z}$, and for a unitary S let $E_k(S)$ be its spectral projection on $I_k(t)$. If S_A and T_B are unitaries of the two parties with $\|(S_A-T_B)\Psi\| \leq h$, then

$$\mathbb{E}_t \sum_k \|(E_k(S_A) - E_k(T_B))\Psi\|^2 \leq 2qh. \tag{15.1}$$

Indeed, for the joint spectral measure ν of (S_A, T_B) on Ψ the sum is twice the ν-probability that the two spectral values z, w lie in different arcs; two points at angular distance d are separated with probability $\min(1, qd) \leq q|z-w|/4$; and $\int |z-w|\,d\nu \leq h$ by Cauchy–Schwarz. Relabel the arcs by j through $k = pj \bmod q$, a bijection, and let P_j^A, P_j^B be the relabelled spectral projections of V_A, V_B. By (15.1), $\mathbb{E}_t \sum_j \|(P_j^A - P_j^B)\Psi\|^2 \leq 2q\delta$.

Conjugation powers. Put $C_a = U_A^{-a}V_AU_A^a$ and $b_a = \|(C_a - \lambda^a V_A)\Psi\|$. For an Alice operator R with $\|R\|_{\mathrm{op}} \leq 2$, $\|RU_A\Psi\| \leq \|R\Psi\| + 2\delta$ by Lemma 2.11(b) with the parties exchanged. Since $C_{a+1} - \lambda^{a+1}V_A = U_A^{-1}(C_a - \lambda^a V_A)U_A + \lambda^a(U_A^*V_AU_A - \lambda V_A)$, this gives $b_1 \leq \delta$ and $b_{a+1} \leq b_a + 3\delta$. With $|\lambda^a - \zeta^a| \leq a\Delta$ and the V mirror,

$$\|(C_a - \zeta^a V_B)\Psi\| \leq (3a+1)\delta + a\Delta \qquad (a \geq 0).$$

The relabelled spectral projections of C_a are $U_A^{-a}P_j^AU_A^a$, and those of ζ^aV_B are P_{j-a}^B, since ζ^a rotates the circle by pa/q. So (15.1) gives

$$\mathbb{E}_t \sum_j \|(U_A^{-a}P_j^AU_A^a - P_{j-a}^B)\Psi\|^2 \leq 2q\big[(3a+1)\delta + a\Delta\big], \tag{15.2}$$

and Lemma 2.11(c) gives $\|(U_B^a - U_A^a)\Psi\| \le a\delta$. No same-party difference of powers is used.

The instruments. Let r be uniform in $\mathbb{Z}/q\mathbb{Z}$, independent of t. Call the mN labels $j = r + bN + a \bmod q$, $0 \le b < m$, $0 \le a < N$, good, and the other $q - mN$ the tail. Alice applies the local instrument whose success branch is

$$W_A = \sum_{b,a} |b,a\rangle \otimes U_A^{-a} P^A_{r+bN+a}, \qquad W_A^* W_A = \sum_{j \text{ good}} P_j^A,$$

and Bob the isometry $W_B = \sum_{b,a} |b,a\rangle \otimes U_B^a P^B_{r+bN+a} + \sum_{j \text{ tail}} |\text{tail}, j\rangle \otimes P_j^B$. Alice announces r, t and her success flag; Bob does not postselect. Averaged over r every label is good with probability mN/q, so the success probability on the original state is exactly $P = mN/q$, with no assumption on the spectrum.

The comparison. Work in the Hilbert direct sum over (r,t) with their uniform weights, the record of (r,t) kept in the purification; for finite-dimensional operators the projections are piecewise constant in t with finitely many breakpoints, so the average over t is a finite weighted sum. Let $\omega = (W_A \otimes W_B \otimes I_E)\Psi$ and

$$y = \sum_{b,a} |b,a\rangle_A |b,a\rangle_B \otimes P^B_{r+bN}\Psi = \Phi_N \otimes \chi, \qquad \chi = \sqrt{N} \sum_b |b,b\rangle \otimes P^B_{r+bN}\Psi,$$

with Φ_N the maximally entangled state of the two a-registers. Both have squared norm P; for y, average $N \sum_b \|P^B_{r+bN}\Psi\|^2$ over r. We make three replacements. (i) Replace ω by the vector with diagonal labels and coefficients $U_A^{-a} P_j^A U_B^a \Psi$, $j = r + bN + a$. For each good j the changed and discarded Bob components have squared norm $2\|P_j^A(I - P_j^B)\Psi\|^2$, and the sum of this over all j is $\sum_j \|(P_j^A - P_j^B)\Psi\|^2$, so the averaged squared error is at most $2q\delta$. (ii) Replace U_B^a by U_A^a in these coefficients. For fixed a the projections P^A_{r+bN+a} are orthogonal, so the squared error is at most $\sum_{a<N} a^2\delta^2 \le N^3\delta^2$. (iii) Replace $U_A^{-a} P^A_{r+bN+a} U_A^a \Psi$ by $P^B_{r+bN}\Psi$. Averaged over r, for each a every label occurs with probability $m/q \le 1/N$, so by (15.2) the squared error is at most $\frac{1}{N}\sum_{a<N} 2q[(3a+1)\delta + a\Delta] \le 3qN\delta + qN\Delta$. By Minkowski's inequality $\|\omega - y\| \le \eta$. The normalized successful output $\omega/\sqrt{P}$ is therefore within $\eta/\sqrt{P}$ of the unit vector $\Phi_N \otimes \chi/\sqrt{P}$, and its two a-registers have singlet fraction at least $1 - \eta^2/P$; the environment, the original systems, the b-registers and the public record are all in the junk.

Rates. If $\gamma = \eta^2/P < \frac{1}{2}$, Alice's heralded filtering (Section 2.8) and (2.4) give $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge P[(1-2\gamma)\log N - h(\gamma)]$, with all communication from Alice to Bob. The decoder depends on the strategy and is not efficient.

Consequences. For a convergent p_n/q_n of $\theta = \arg\lambda/2\pi$, $|\theta - p_n/q_n| < 1/(q_n q_{n+1})$ gives $q_n^2|\lambda - e^{2\pi i p_n/q_n}| \le 2\pi$. Take $N = \lfloor cq_n \rfloor$ with a small fixed c and $\delta \le q_n^{-4}$: then $P \ge 1 - c$, the first two terms of η tend to zero and $qN\Delta \le 2\pi c$, so $\gamma = O(c)$ for large n and the rate is at least $0.99\log q_n - O(1)$ once c is small. With $\varepsilon_n = C_G^{-2} q_n^{-8}$ every strategy of loss at most ε_n has $\delta \le q_n^{-4}$, which gives the logarithmic bound along $\varepsilon_n \to 0$. If $q_{n+1} \le K q_n^B$, given δ choose the largest $q_n \le \delta^{-1/4}$; then $q_n \ge K^{-1/B}\delta^{-1/(4B)}$, and the bound holds at every small ε. For algebraic non-torsion λ, Baker–Wüstholz [6] gives $|\lambda^q - 1| \ge aq^{-A}$ for $q \ge 1$, and $|\lambda^{q_n} - 1| \le 2\pi/q_{n+1}$ yields $q_{n+1} \le (2\pi/a) q_n^A$. For the gadget, the admitted set is semialgebraic over the real algebraic numbers by Tarski–Seidenberg [129], so it has finitely many connected components, points or arcs; a nontrivial arc contains a root of unity, so only finitely many isolated points remain, and these have real algebraic coordinates. □

These statements do not cover an uncontrolled central-operator phase or unrestricted noncommutative encodings, and they provide no fixed rational game compiler; no game below the logarithmic floor is obtained.

15.8 Finite-support variants give integer growth

Proposition 15.9 (Finite supports)**.** *Let H_3 act transitively on a set of objects. Fix a k-element set $S \subseteq Y$ and suppose that the stabilizer K of a base object o satisfies $\mathrm{Fix}_{H_3}(S) \leq K \leq \mathrm{Stab}_{H_3}(S)$, the pointwise and the setwise stabilizer of S. Then $L = K/\mathrm{Fix}_{H_3}(S)$ is a subgroup of S_k, the orbit is $\mathrm{Inj}([k], Y)/L$, the object ho has support $h(S)$, and the number of objects supported in the first N points is $|\mathrm{Inj}([k], [N])|/|L|$, of integer polynomial order N^k.*

Proof. Fix a bijection $\iota : [k] \to S$. H_3 contains all finitary permutations, so it acts transitively on injections $[k] \to Y$ by $h \cdot f = h \circ f$, with stabilizer $\mathrm{Fix}(S)$ at ι, and restriction to S maps $\mathrm{Stab}(S)$ onto $\mathrm{Sym}(S) \cong S_k$ with kernel $\mathrm{Fix}(S)$. So $\mathrm{Fix}(S)$ is normal in K and $L = K/\mathrm{Fix}(S)$ embeds in S_k, acting on injections by precomposition. The map $hK \mapsto (h \circ \iota)L$ is well defined, since $h\kappa \circ \iota = h \circ \iota \circ \iota^{-1}\kappa|_S\iota$ for $\kappa \in K$, and bijective; and L acts freely on injections. The count follows. □

The two-sided hypothesis is needed: the lower bound $\mathrm{Fix}(S) \leq K$ alone does not make S a least support. A trivial one-object action has $K = H_3$, which contains $\mathrm{Fix}(S)$ for every finite S, so it admits declared supports of every size, while its orbit has one element.

Finite colours, orderings and finite quotient patterns within this architecture therefore do not create a fractional exact law; Theorem 14.9 escapes by replication, not by refuting the premise. This is standard finite-support reasoning, not a classification of all finitely presented actions.

15.9 Operator accuracy without trace visibility

Proposition 15.10 (Invisible witnesses)**.** (a) *A D-dimensional unitary W with $\mathrm{rank}(W - I) \leq r$ has $\|W - I\|_{2,D}^2 \leq 4r/D$.* (b) *If J lies in the normal closure of a word w in a fixed extension presentation, then in the tracial normalized setting with root relator error e, $\|J - I\|_2 \leq C\|w - I\|_2 + C'e$.* (c) *Consider the relations $S^*S = I - P_N$, $SS^* = I - P_0$, $XS = SX^2$, $XP_0 = \frac{1}{2}P_0$, $XP_N = 0$, with P_0, P_N projections and X positive. For every D they have matrix models with all relations exact except $\|XP_N\| = 2^{-2^{D-1}}$, and with $\|P_0\| = 1$ but $\mathrm{tr}_D(P_0) = 1/D$; and every exact model in a finite von Neumann algebra with a faithful normal tracial state, in particular every exact finite-dimensional model, has $P_0 = 0$.*

Proof. (a) $\mathrm{Tr}(W - I)^*(W - I) \leq 4\,\mathrm{rank}(W - I)$. (b) A fixed derivation writes J as a product of conjugates of $w^{\pm 1}$ and relators; unitary invariance of the normalized norm and telescoping give the bound. (c) The shift $Se_i = e_{i+1}$ $(i < D - 1)$, $Se_{D-1} = 0$ on $\mathbb{C}^D$, with P_0 and P_N the projections onto e_0 and e_{D-1} and $X = \mathrm{diag}(2^{-2^i})$, satisfies all but the last relation exactly. In an exact model, let $P_0 v = v \neq 0$. Then $Xv = \frac{1}{2}v$, and $XS = SX^2$ gives $XS^k v = 2^{-2^k} S^k v$ for every k by induction. Since X is self-adjoint, each $S^k v$ with $S^k v \neq 0$ lies in an eigenspace with nonzero eigenvalue, hence is orthogonal to $\ker X \supseteq \mathrm{ran}\, P_N$, so $\|S^{k+1}v\|^2 = \langle S^k v, (I - P_N)S^k v\rangle = \|S^k v\|^2$. Thus S^k is isometric on $\mathrm{ran}\, P_0$, the projections $E_k = S^k P_0 S^{*k}$ are equivalent to P_0, and they lie in distinct eigenspaces of X, hence are mutually orthogonal. A faithful normal tracial state τ then gives $(n+1)\tau(P_0) = \tau(E_0 + \cdots + E_n) \leq 1$ for every n, so $\tau(P_0) = 0$ and $P_0 = 0$. For finite matrices alone positivity of X is not needed: $\ker S = \ker S^*S = \mathrm{ran}\, P_N \subseteq \ker X$, so inductively every $S^k v$ is a nonzero eigenvector of X with eigenvalue 2^{-2^k}; these are linearly independent, which is impossible in finite dimension. □

So exponentially accurate operator relations with a bounded-rank witness have normalized visibility $O(D^{-1/2})$, and normal-closure compilation cannot promote vanishing trace visibility into $J = -I$ at vanishing error on the whole space. Baumslag–Solitar near-collision

models [134, 62] have exponentially accurate operator relations of exactly this invisible kind. A central marker that is visible in operator norm can always be compressed to a negative sector without loss (Lemma 17.7); for these models the obstruction then reappears on the compressed sector (Propositions F.1 and F.3). The known weak dimension lower bound for $\mathrm{BS}(2,3)$ is not a polylogarithmic upper bound, and no claim is made that these routes are impossible in general.

16 A Bell-gap contrast and finite blocks

This section uses the Bell-gap contract of Section 2.8 for its conditional I_{3322} statement, then returns to classes (S) and (C) for the finite-block observation. The two contracts are compared in Section 1.6; whether Theorem A extends beyond classes (S) and (C) is open, and this section does not decide it.

16.1 The I_{3322} contrast

We use [53] in its arXiv version 2, whose mathematical content is that of version 4.1.0 of its archive. In its Collins–Gisin normalization it proves finite-dimensional nonattainment of the quantum supremum S, certified to lie in $(0.2508753845015185, 0.250875388108398]$, for mixed states and POVMs [53, Theorem 1]; attainment by a normal spatial strategy on $\ell^2(\mathbb{Z}) \otimes \ell^2(\mathbb{Z})$ [53, Theorem 2]; and $D(\varepsilon) = \Theta(\ln(1/\varepsilon))$ for the least local dimension at Bell gap $\varepsilon = S - B$, with eventual upper coefficient 23.9010650 in natural logarithms, not asserted optimal and with no explicit onset [53, Theorem 6]. Nonattainment was proved independently by Pauwels [119, Theorem 2], and nonattainment together with attainment by a spatial strategy on separable infinite-dimensional Hilbert spaces by Coladangelo [37, Theorem 11.1]; the infinite-dimensional construction of the optimum, and the conjecture that finite dimension does not reach it, are due to Pál and Vértesi [116]. The tail estimate we use is that of [53, Section 5.2]: the Schmidt coefficients of the spatial attainer constructed there decay exponentially at both ends. Machine checks cover only parts of these works. Pauwels' Lean formalization covers pure states and projective measurements, and the reduction from mixed states and POVMs lies outside it [119, Appendix D]; the Lean checks of [53] cover selected numerical, scalar and finite accounting statements, and its analytic proofs, including spatial attainment and the tail estimate, are not formalized. We carry that status.

Theorem 16.1 (A conditional Bell-gap contrast)**.** *Conditional on the spatial attainment theorem [53, Theorem 2] and the tail bound $\kappa_\pm \geq 0.08367827985$ of [53, Section 5.2], the infimum of the original one-way distillable entanglement, in either direction, over strategies with Bell gap at most ε is $O(1)$ as $\varepsilon \to 0$, although no finite-dimensional strategy attains S and the least dimension grows like $\ln(1/\varepsilon)$.*

Proof. The normal pure attainer of [53] is $\sum_j \lambda_j |j, j\rangle$ with geometric tails: for each small $\eta > 0$, $\lambda_j^2 \leq A_\eta \exp[-(\kappa_\pm - \eta)|j|]$ on the two tails, with $\kappa_\pm \geq 0.08367827985$ per index [53, Section 5.2]. So $p_j = \lambda_j^2$ has finite Shannon entropy $H(p)$; finite norm alone would not give this. The near-optimal finite strategies of [53, Section 5.2] are the renormalized windows of this same Schmidt vector on finite intervals I, with $Z_I = \sum_{j \in I} p_j \to 1$, and their Bell gaps tend to zero. Their Schmidt distributions $q_j = p_j / Z_I$ satisfy

$$H(q) = Z_I^{-1} \sum_{j \in I} (-p_j \log p_j) + \log Z_I \leq H(p)/Z_I,$$

since every term is nonnegative and $\log Z_I \leq 0$. These states are pure, so both directional rates equal $H(q)$, which stays bounded. □

No matching positive lower bound, exact limit or minimizing state is asserted, and the logarithmic dimension theorem does not force any growth of E_{D}. The contract cannot be changed by rescaling: the nonsignaling value of I_{3322} is one, above the quantum optimum $S \simeq 0.2509$. With the joint coefficients $\left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \end{smallmatrix}\right)$ and marginal coefficients $(-1, 0, 0)$ and $(-2, -1, 0)$, the behaviour with all marginals $\frac{1}{2}$ and $P(1,1) = \frac{1}{2}$ on the positive entries and 0 on the negative ones has value $-2 + \frac{6}{2} = 1$, so an affine rescaling $aB + b$ with quantum maximum one would score $1 + a(1 - S) > 1$ on it, impossible for a probability of winning.

16.2 Normal perfect strategies have finite perfect blocks

Remark 16.2 (Finite blocks; known). In either class of Definition 7.1, any normal spatial perfect strategy, even with infinite-dimensional ambient systems and an arbitrary normal mixed input, yields a finite-dimensional perfect strategy.

Proof. For the synchronous class this is the theorem of Kim, Paulsen and Schafhauser that $C_q^s(n,m) = C_{qs}^s(n,m)$ [83, Theorem 3.10 and Corollary 3.11]. Directly: after local projectivization, perfect diagonal consistency gives $P_{xa}^A \Psi = Q_{xa}^B \Psi$ on a purification; taking partial traces, the trace-class Bob marginal commutes with every Bob projection. A nonzero eigenvalue has a finite-dimensional eigenspace invariant under these projections, and the losing products, which annihilate the support (move one factor to Alice), vanish on it; the transpose maximally entangled realization of this block is a finite perfect strategy. For a native constraint-variable game the same argument applies with the incidence mirrors and Alice's commuting legal row projections, which is the standard operator-solution reconstruction [34]. □

So a game in the classes of Theorem A without a finite perfect strategy has no normal perfect strategy at all, of any entropy. Normal infinite-dimensional correlations that have no finite-dimensional realization were constructed by Coladangelo and Stark [39]; the finite-block conclusion has a more restricted scope, perfect strategies of games in the two classes. The I_{3322} attainer is normal and of finite entropy, so a normality-preserving adapter that turned it into a perfect strategy of a fixed target game in the classes of Theorem A without a finite-dimensional perfect strategy would give that target a normal spatial perfect strategy, which it does not have. For such a target, any sequence of finite-dimensional near-perfect strategies has diverging total one-way distillable entanglement, by Theorem A. A claim about the additional resources an adapter needs would require a specified conversion contract, which we do not give.

17 The logarithmic floor: a class theorem and a criterion

Theorem A gives, for each target r and fidelity $1 - \gamma$, a loss threshold $\varepsilon_0(G, r, \gamma)$, but no rate. Problem 18.2 asks whether every game of Theorem A has at least logarithmic demand. This section proves the floor on an explicit class: games that recover, with actual mirrors on the original state, a sign relation and an even dilation (Theorem 17.3), with a game compiled from Abels' finitely presented solvable group as an instance (Corollary 17.6). It then proves a sufficient criterion for a counterexample (Proposition 17.8), through an operator-norm compression of a central involution (Lemma 17.7). The candidate classes that the criterion excludes, and the obstructions to proving the floor in general by the method of Theorem A, are in Appendix F.

17.1 The sign-and-dilation class

Definition 17.1 (The sign-and-dilation interface)**.** Let $m \geq 2$ be an even integer and $C_G > 0$. A finite game G has the *sign-and-dilation interface* with constant C_G if every finite-dimensional mixed strategy of loss ε, after local projective dilation, with Ψ a purification of the dilated state, has actual local unitaries X_P, Y_P, W_P on each party $P \in \{A, B\}$ such that, with $e = C_G\sqrt{\varepsilon}$,

(I1) $\|(U_A - U_B)\Psi\| \leq e$ and $\|(U_A^* - U_B^*)\Psi\| \leq e$ for $U = X, Y, W$;

(I2) $\|(X_B Y_B X_B^* - Y_B^m)\Psi\| \leq e$ and $\|(X_A^* Y_A X_A - Y_A^m)\Psi\| \leq e$;

(I3) $\|(W_P Y_P + Y_P W_P)\Psi\| \leq e$ for $P = A, B$.

The orientation in (I2) is that of Proposition 6.1: Alice realizes the opposite algebra. The sign witness W need not be an involution. No access to the environment, to a pure-state decomposition or to Bob's Schmidt partners is assumed, and Definition 17.1 is a property of a game, checkable from its algebra: Corollary 17.6 derives it from relations $[w, y] = J$ and $xyx^{-1} = y^m$ in a source compiled by Theorem 5.6. In the negative quotient, $[w, y] = J = -1$ says $wy = -yw$.

The first step turns the unitary sign witness into a reflection, by a random semicircle cut applied to commuting operators of the two parties. For $\vartheta \in [0, 1)$ and z on the unit circle let $h_\vartheta(z) = 1$ if the angle of z, in turns, lies in $[\vartheta, \vartheta + \frac{1}{2})$ modulo one, and $h_\vartheta(z) = -1$ otherwise; then $h_\vartheta(-z) = -h_\vartheta(z)$, and $h_\vartheta(U)$ is a reflection for every unitary U, with $h_\vartheta(VUV^*) = V h_\vartheta(U) V^*$.

Lemma 17.2 (Cuts across the parties)**.** *Let U_A and V_B be unitaries of the two parties, so that they commute, and Ψ a unit vector. For ϑ uniform in $[0, 1)$,*

$$\mathbb{E}_\vartheta \|(h_\vartheta(U_A) - h_\vartheta(V_B))\Psi\|^2 \;\leq\; 2\,\|(U_A - V_B)\Psi\|.$$

Proof. Let ν be the joint spectral measure of the commuting normal operators U_A, V_B in the vector Ψ, a probability measure on pairs (z, w) of points of the circle. Then $\|(h_\vartheta(U_A) - h_\vartheta(V_B))\Psi\|^2 = 4\,\nu\{h_\vartheta(z) \neq h_\vartheta(w)\}$. The two endpoints of the semicircle are half a turn apart, so for z, w at circle distance $d \leq \frac{1}{2}$ in turns at most one of them lies on the shorter arc between z and w, and the cut separates z from w with probability $2d$. As $d \leq \frac{1}{4}|z - w|$,

$$\begin{aligned}\mathbb{E}_\vartheta \|(h_\vartheta(U_A) - h_\vartheta(V_B))\Psi\|^2 &= 8 \int d\,d\nu \\ &\leq 2 \int |z - w|\,d\nu \leq 2\Big(\int |z - w|^2\,d\nu\Big)^{1/2} = 2\|(U_A - V_B)\Psi\|.\end{aligned}$$

□

Theorem 17.3 (A logarithmic floor on the sign-and-dilation class)**.** *Let G have the sign-and-dilation interface for an even $m \geq 2$. There are $\varepsilon_G > 0$ and $C > 0$ such that every finite-dimensional mixed strategy of loss $\varepsilon < \varepsilon_G$ satisfies*

$$E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \;\geq\; \frac{\log(1/\varepsilon)}{4\log m} - C\log\log(1/\varepsilon),$$

and hence $d_G^{\rightarrow}(\varepsilon) \geq \log(1/\varepsilon)/(8\log m) - O(1)$ for $0 < \varepsilon < 1$. Moreover there is $c_G > 0$ such that, for every $r \geq 1$ and $0 < \gamma < 1$, every finite-dimensional mixed strategy of loss at most $c_G\gamma^8 r^{-64} m^{-4r}$ admits product local isometries whose output is within $\sqrt{\gamma}$, in vector norm on any purification, of $\Phi_{2^r} \otimes \chi$, the environment included in the junk χ. So in this class the threshold of Theorem A can be taken at least $\exp(-C_{G,\gamma} r)$.

Proof. Fix a strategy of loss ε, dilate it locally, and let X_P, Y_P, W_P and $e = C_G\sqrt{\varepsilon}$ be as in Definition 17.1; assume $e \le 1$. Directional rates are unchanged by the local isometries of the dilation.

Step 1: a reflection for the sign. Put $T_P = h_\vartheta(W_P)$, with one cut ϑ for both parties. By Lemma 17.2 and (I1), $\mathbb{E}\|(T_A - T_B)\Psi\|^2 \le 2e$. On one party P, with $\bar{P}$ the other, put $D = W_P Y_P + Y_P W_P$, of operator norm at most two. Then $Y_P W_P Y_P^* + W_P = D Y_P^*$, and moving Y_P^* through D to the other party costs $2e$ by (I1), so $\|(Y_P W_P Y_P^* + W_P)\Psi\| \le \|D\Psi\| + 2e \le 3e$ by (I3); with the mirror of W, $\|(Y_P W_P Y_P^* - (-W_{\bar{P}}))\Psi\| \le 4e$. The unitaries $Y_P W_P Y_P^*$ and $-W_{\bar{P}}$ act on different parties, so Lemma 17.2 and $h_\vartheta(Y_P W_P Y_P^*) = Y_P T_P Y_P^*$, $h_\vartheta(-W_{\bar{P}}) = -T_{\bar{P}}$ give $\mathbb{E}\|(Y_P T_P Y_P^* + T_{\bar{P}})\Psi\|^2 \le 8e$. Adding the mirror of T, $\mathbb{E}\|(Y_P T_P Y_P^* + T_P)\Psi\|^2 \le 2 \cdot 8e + 2 \cdot 2e = 20e$. Finally $T_P Y_P T_P + Y_P = T_P(Y_P T_P Y_P^* + T_P) Y_P$, and moving the right factor Y_P to the other party costs at most $2e$, so $\mathbb{E}\|(T_P Y_P T_P + Y_P)\Psi\|^2 \le 2 \cdot 20e + 2 \cdot 4e^2 \le 48e$. The sum of the two squared sign errors and the squared mirror error of T has expectation at most $98e$, so some cut has all three at most $\sqrt{98e} < 10\sqrt{e}$. The cut depends on the strategy; it is a choice of local operators, not a postselection. With it, the actual reflections T_P satisfy $\|(T_A - T_B)\Psi\| \le 10\sqrt{e}$ and $\|(T_P Y_P T_P + Y_P)\Psi\| \le 10\sqrt{e}$ on both parties, the unitaries X_P, Y_P satisfy (I1) and (I2) with $e \le \sqrt{e}$, and by Lemma 2.11(d) all square-root-state commutators of T_P, X_P, Y_P are at most $10\sqrt{2}\sqrt{e}$. Put $\delta = 15\sqrt{e}$, which bounds all of these.

Step 2: growing extraction. The data of Step 1 are those of one copy in the proof of Theorem 11.9, with $q = 2$, the reflection T in place of the rotation control and δ in place of e: $XYX^* = Y^m$ and $TYT = -Y$ on Bob, the opposite relations on Alice, and all mirrors, inverse mirrors and square-root commutators at most δ. For $q = 2$ that proof uses only that m is even (so that m^{l-j} is even for $j < l$), and for one copy it uses no cross-copy relation. Its steps *powers without an exponential square*, *common roots and controls*, *spectral cuts* and *extraction*, with the single-copy bound there, therefore give, for every integer $L \ge 1$ with $C_m m^L \delta \le 1$, product local isometries whose output is within

$$\gamma_L \;\le\; C_m L^4 (m^L \delta)^{1/4}$$

of $\Phi_{2^L} \otimes \chi$, the environment included in χ, where C_m depends only on m.

Step 3: the rate. Put $x = \log(1/\delta)$ and $L = \lfloor (x - 24\log(2+x))/\log m \rfloor$. For ε small, $L \ge 1$, and $m^L \delta \le (2+x)^{-24}$, so $C_m m^L \delta \le 1$ and $\gamma_L \le C_m L^4 (2+x)^{-6} \le C'_m (2+x)^{-2}$, since $L \le x$. The output has singlet fraction at least $1 - \gamma_L^2$, and (2.4) gives $E_{\mathrm{D}}^{A\to B}(\rho_{AB}) \ge (1 - 2\gamma_L^2)L - h(\gamma_L^2) = L - o(1)$, because $L\gamma_L^2 \le C(2+x)^{-3}$. Finally $L \ge x/\log m - 24\log(2+x)/\log m - 1$ and $x = \frac{1}{2}\log(1/e) - \log 15 = \frac{1}{4}\log(1/\varepsilon) - \frac{1}{2}\log C_G - \log 15$. This is the first bound. For small ε it implies the second, and the second holds trivially, by enlarging its constant, on the remaining range $\varepsilon_G \le \varepsilon < 1$ and wherever $L < 1$.

The threshold. Take $L = r$ in Step 2. With $\delta = 15 C_G^{1/2} \varepsilon^{1/4}$ we get $\gamma_r \le C_{G,m} r^4 m^{r/4} \varepsilon^{1/16}$, and $\gamma_r \le \sqrt{\gamma}$ as soon as $\varepsilon \le c_G \gamma^8 r^{-64} m^{-4r}$ for a small c_G, which also ensures $C_m m^r \delta \le 1$ and $e \le 1$. Since $\gamma < 1$ and $r \ge 1$, this threshold is at least $\exp(-C_{G,\gamma} r)$. □

All operations act on Alice and Bob, and the bound concerns the original mixed state; no marginal entropy, entanglement of formation or replacement strategy is used. The extraction architecture is that of Theorem 11.9; what is new here is that the sign witness may be an arbitrary unitary, rounded across the parties, and that the class is defined by an interface rather than by a particular source. The coefficient is not claimed optimal.

17.2 An instance: Abels' group

Let $R = \mathbb{Z}[\frac{1}{2}]$ and let $A_4(R)$ be the group of upper triangular 4×4 matrices over R with diagonal $(1, t_{22}, t_{33}, 1)$, $t_{22}, t_{33} \in R^\times$, in the notation of Cornulier and Tessera [43]. Abels

proved that $A_4(\mathbb{Z}[1/p])$ is finitely presented [1]; see [43, Introduction]. The elementary matrices $e_{14}(r) = I + rE_{14}$ form a central subgroup, since every matrix of $A_4(R)$ has first column e_1 and last row e_4^T. Put

$$B = A_4(R)/\langle e_{14}(1)\rangle, \qquad w = e_{12}(1), \quad y = e_{24}(\tfrac{1}{2}), \quad x = \mathrm{diag}(1,2,1,1), \quad J = e_{14}(\tfrac{1}{2}).$$

The quotient by the cyclic central subgroup generated by one element adds one relator, so B is finitely presented.

Lemma 17.4. *In B, J is a central involution, $J \neq 1$, $[w,y] = J$ and $xyx^{-1} = y^2$. Every finite-dimensional unitary representation of B sends J to I.*

Proof. Matrix multiplication gives $[e_{12}(s), e_{24}(t)] = e_{14}(st)$, so $[w,y] = e_{14}(\frac{1}{2}) = J$; and conjugation by x multiplies the entry $(2,4)$ by $t_{22}/t_{44} = 2$, so $xyx^{-1} = e_{24}(1) = y^2$. J is central, $J^2 = e_{14}(1) = 1$ in B, and $J \neq 1$ because $\frac{1}{2} \notin \mathbb{Z}$. If a finite-dimensional unitary representation π had $\pi(J) \neq I$, its (-1)-eigenspace, invariant because J is central, would carry unitaries $A = \pi(w)$, $X = \pi(x)$, $Y = \pi(y)$ with $AYA^* = -Y$ and $XYX^* = Y^2$ exactly, in some dimension $k \geq 1$; Proposition 17.5 below with $\alpha = \beta = 0$ excludes this. □

Proposition 17.5 (A dyadic bound)**.** *If unitaries A, X, Y on $\mathbb{C}^k$ satisfy $\|AYA^* + Y\|_{2,k} \leq \alpha$ and $\|XYX^* - Y^2\|_{2,k} \leq \beta$ in the normalized Hilbert–Schmidt norm, then $k \geq 1/(\alpha+\beta)$. In particular the bound holds with operator norms.*

Proof. Apply Lemmas 9.1 and 9.2 with the flat state $\rho = I/k$, for which $\|P\xi - \xi Q\|_2 = \|P - Q\|_{2,k}$ and conjugate unitaries have the same spectral measure μ. With (9.2), $W_1(S_*\mu, \mu) = W_1(\mu_{-Y}, \mu_{AYA^*}) \leq \frac{\pi}{2}\alpha$ and $W_1(D_*\mu, \mu) = W_1(\mu_{Y^2}, \mu_{XYX^*}) \leq \frac{\pi}{2}\beta$, so $W_1(\mu, h) \leq \frac{\pi}{2}(\alpha+\beta)$ by Lemma 9.1(c), while $W_1(\mu, h) \geq \pi/(2k)$ by Lemma 9.1(d), μ having at most k atoms. □

Fix a finite generating set S of B containing w, x, y, J and a finite presentation $\langle S \mid R_B\rangle$ of B that contains the relators J^2, $[J, s]$ for $s \in S$, $wyw^{-1}y^{-1}J^{-1}$ and $xyx^{-1}y^{-2}$; adding valid relators to a finite presentation keeps it a presentation of B. Let G_B be the Boolean constraint game of Theorem 5.6 for this presentation and the marked generator J, with its uniform question law.

Corollary 17.6 (The Abels game)**.** *G_B is a game in class (C) of quantum value one with no finite-dimensional perfect strategy, every incidence is tested, and it has the sign-and-dilation interface with $m = 2$. Hence*

$$d^{\rightarrow}_{G_B}(\varepsilon) \;\geq\; \tfrac{1}{4}\log(1/\varepsilon) - O(\log\log(1/\varepsilon)), \qquad \textit{in particular} \qquad d^{\rightarrow}_{G_B}(\varepsilon) \geq \tfrac{1}{8}\log(1/\varepsilon) - O(1).$$

Proof. No finite perfect strategy. A finite-dimensional perfect strategy of G_B gives a nonzero finite-dimensional representation of $\mathcal{A}(B_G)$ (Section 7.1), and composing with the unital $*$-homomorphism $F : A_B^- \to \mathcal{A}(B_G)$ of Theorem 5.6 gives a unitary representation of B with $J = -I$ in finite dimension, which Lemma 17.4 excludes.

Value one. This is a chain of seven steps. (1) B is solvable, hence amenable, and countable, so its reduced group C^*-algebra $C_r^*(B)$ is quasidiagonal by the theorem of Tikuisis, White and Winter [137, Corollary C]. (2) By Voiculescu's characterization, which [137, Definition 1.1] takes as the definition, there are unital completely positive maps $\phi_n : C_r^*(B) \to M_{D_n}$ with $\|\phi_n(ab) - \phi_n(a)\phi_n(b)\| \to 0$ and $\|\phi_n(a)\| \to \|a\|$ for all a, b in any fixed finite set; we use the finite set of the left translations $\lambda_s^{\pm 1}$, $s \in S$, and their products along the prefixes of the relators of R_B. (3) Let $U_{n,s}$ be a unitary polar part of $\phi_n(\lambda_s)$. Since ϕ_n is unital and $*$-preserving, $\|\phi_n(\lambda_s)^*\phi_n(\lambda_s) - I\| = \|\phi_n(\lambda_s^*)\phi_n(\lambda_s) - \phi_n(\lambda_s^*\lambda_s)\| \to 0$, and the scalar inequality $|1-t| \leq |1-t^2|$ gives $\|U_{n,s} - \phi_n(\lambda_s)\| \leq \|\phi_n(\lambda_s)^*\phi_n(\lambda_s) - I\|$. Telescoping each relator r against the product of the $\phi_n(\lambda_s^{\pm 1})$, which are contractions, and then against

$\phi_n(\lambda_r) = \phi_n(1) = I$ by asymptotic multiplicativity, gives $\eta_n = \max_{r \in R_B} \|r(U_n) - I\| \to 0$. (4) Since $J \neq 1$ and λ is faithful, λ_J is a self-adjoint unitary other than 1, so $\|1 - \lambda_J\| = 2$, and $\|U_{n,J} - I\| \geq \|\phi_n(1 - \lambda_J)\| - \|U_{n,J} - \phi_n(\lambda_J)\| \to 2$. (5) For large n, $\|U_{n,J} - I\| > \eta_n$, and the compression lemma (Lemma 17.7) gives unitaries $V_{n,s}$ on some $\mathbb{C}^{k_n}$, $k_n \geq 1$, with $V_{n,J} = -I$ and $\max_r \|r(V_n) - I\| \leq (1 + 3L_B)\eta_n$, where L_B is the largest relator length. (6) With the flat state, these are models of A_B^- at precision $(1 + 3L_B)\eta_n$: the normalized Hilbert–Schmidt norm is at most the operator norm, state commutators vanish and $V_{n,J} + I = 0$. So the flat profile $h_B(e)$ is finite for every $e > 0$. (7) The value-one clause of Theorem 5.6 gives $\omega_q(G_B) = 1$.

The interface. Proposition 6.1, applied to the lower map of Theorem 5.6, gives, for $e = C\sqrt{\varepsilon}$, Bob unitaries U_s and Alice unitaries U_s^A with mirrors and inverse mirrors at most e, every relator of R_B with residual at most e on Bob and its reversed word with residual at most e on Alice, and $\|(U_J + I)\Psi\| \leq e$. Put $X_P = U_x^P$, $Y_P = U_y^P$ and $W_P = U_w^P$; (I1) holds. For (I2), let $F = U_x U_y U_x^* U_y^{-2}$, the Bob value of the relator $xyx^{-1}y^{-2}$; then $X_B Y_B X_B^* - Y_B^2 = (F - I)Y_B^2$, and moving Y_B^2 to Alice costs $2 \cdot 2e$ by Lemma 2.11(c), so

$$\|(X_B Y_B X_B^* - Y_B^2)\Psi\| \leq \|(F - I)\Psi\| + 2 \cdot 2e \leq (2m + 1)e = 5e.$$

On Alice the reversed relator is $Y_A^{-2} X_A^* Y_A X_A$, and $X_A^* Y_A X_A - Y_A^2 = Y_A^2(Y_A^{-2} X_A^* Y_A X_A - I)$ has residual at most e. For (I3), let $F' = W_B Y_B W_B^* Y_B^* U_J^*$, the Bob value of $wyw^{-1}y^{-1}J^{-1}$. Then $W_B Y_B + Y_B W_B = (F' - I)U_J Y_B W_B + (U_J + I)Y_B W_B$; moving the three-letter word $U_J Y_B W_B$ to Alice costs $2 \cdot 3e$, and the two-letter word $Y_B W_B$ costs $2 \cdot 2e$, so $\|(W_B Y_B + Y_B W_B)\Psi\| \leq (e + 6e) + (e + 4e) = 12e$. On Alice the reversed relator gives $G = U_J^{A*} Y_A^* W_A^* Y_A W_A$ with $\|(G - I)\Psi\| \leq e$, and $Y_A W_A + W_A Y_A = W_A Y_A (U_J^A (G - I) + (U_J^A + I))$, where $\|(U_J^A + I)\Psi\| \leq \|(U_J^A - U_J)\Psi\| + \|(U_J + I)\Psi\| \leq 2e$; so it has residual at most $3e$. Hence G_B has the interface with constant $12C$ and $m = 2$, and Theorem 17.3 gives the bounds. □

The source B is the Abels quotient considered in Appendix F.4; there it appears as a candidate that the criterion below excludes as a counterexample, consistently with the floor proved here. No upper bound or sharp coefficient is claimed for G_B, and no explicit export of the game is supplied; its existence and properties follow from Theorem 5.6.

17.3 Compression and the counterexample criterion

A central involution that is visible in operator norm can be moved to a negative sector without loss of accuracy.

Lemma 17.7 (Operator-norm compression)**.** *Let $\langle S \mid R \rangle$ be a finite presentation with a generator $J \in S$ and the relators J^2 and $[J, s]$ for $s \in S$, and let L be the largest length of a relator. Let unitaries U_s on $\mathbb{C}^D$ satisfy $\max_{r \in R} \|r(U) - I\|_{\mathrm{op}} \leq \eta$ and $\|U_J - I\|_{\mathrm{op}} > \eta$. Then there are a nonzero subspace $\mathbb{C}^k \subseteq \mathbb{C}^D$ and unitaries V_s on it with $V_J = -I_k$ and $\max_{r \in R} \|r(V) - I_k\|_{\mathrm{op}} \leq (1 + 3L)\eta$. Moreover V_s is the restriction of a unitary B_s on $\mathbb{C}^D$ with $\|B_s - U_s\|_{\mathrm{op}} \leq 3\eta$ that commutes with the projection onto $\mathbb{C}^k$.*

Proof. For z on the unit circle let $\mathrm{s}(z) \in \{1, -1\}$ be a nearest sign, with a fixed choice at ties. With $a = |\operatorname{Re} z|$, $|z - \mathrm{s}(z)|^2 = 2(1 - a) \leq 4(1 - a)(1 + a) = |z^2 - 1|^2$. So $H = \mathrm{s}(U_J)$ is a self-adjoint involution with $\|U_J - H\|_{\mathrm{op}} \leq \|U_J^2 - I\|_{\mathrm{op}} \leq \eta$. If $H = I$, then $\|U_J - I\|_{\mathrm{op}} \leq \eta$, which is excluded; so $P = \frac{1}{2}(I - H)$ is a nonzero projection, of rank $k \geq 1$. For $s \neq J$, the relator $[J, s]$ gives $\|HU_s - U_s H\|_{\mathrm{op}} \leq \|U_J U_s - U_s U_J\|_{\mathrm{op}} + 2\|H - U_J\|_{\mathrm{op}} \leq 3\eta$.

Put $A_s = \frac{1}{2}(U_s + HU_s H)$ and $C_s = U_s - A_s$. Relative to H, A_s is block diagonal and C_s block off-diagonal, and $t_s = \|C_s\|_{\mathrm{op}} = \frac{1}{2}\|HU_s - U_s H\|_{\mathrm{op}} \leq \frac{3}{2}\eta$; also $t_s \leq 1$. The block-diagonal part of $U_s^* U_s = I$ is $A_s^* A_s + C_s^* C_s = I$, so every singular value of A_s is at least $\sqrt{1 - t_s^2}$. In

each of the two square diagonal blocks of A_s choose a unitary polar part, which exists even if the block is singular, and let B_s be their direct sum. Then B_s commutes with H, and $\|B_s - A_s\|_{\mathrm{op}} = 1 - \sigma_{\min}(A_s) \le 1 - \sqrt{1 - t_s^2} \le t_s$, so $\|B_s - U_s\|_{\mathrm{op}} \le 2t_s \le 3\eta$. Put $B_J = H$ separately; then $\|B_J - U_J\|_{\mathrm{op}} \le \eta$. Telescoping, with $\|B_s^{-1} - U_s^{-1}\|_{\mathrm{op}} = \|B_s - U_s\|_{\mathrm{op}}$, every relator has $\|r(B) - I\|_{\mathrm{op}} \le \|r(U) - I\|_{\mathrm{op}} + 3|r|\eta \le (1 + 3L)\eta$. All B_s commute with P, so their restrictions V_s to the range of P are unitaries, products and inverses restrict exactly, $V_J = H|_{P\mathbb{C}^D} = -I_k$, and the operator norm does not increase on an invariant subspace. □

No lower bound on k/D is needed, and nothing is lost in passing to the sector. In normalized Hilbert–Schmidt norm the same compression would pay for the trace fraction of P; in operator norm it does not. So a marked central involution that is operator-visible is never hidden by a small trace fraction: Proposition 15.10 concerns witnesses of bounded rank that are not central.

Proposition 17.8 (A criterion for a counterexample to the floor)**.** *Let $\langle S \mid R\rangle$ be a finite presentation with a marked central involution $J \in S$ and the relators J^2, $[J, s]$, such that no finite-dimensional unitary representation sends J to $-I$. Suppose there are unitaries $U_s^{(n)}$ on $\mathbb{C}^{D_n}$ with $\eta_n = \max_r \|r(U^{(n)}) - I\|_{\mathrm{op}} \to 0$, $\|U_J^{(n)} - I\|_{\mathrm{op}} > \eta_n$, and*

$$\log D_n = o\big(\log(1/\eta_n)\big).$$

Then the game G of Theorem 5.6 for this presentation lies in class (C), has quantum value one and no finite-dimensional perfect strategy, and $d_G^{\rightarrow}(\varepsilon_n)/\log(1/\varepsilon_n) \to 0$ along $\varepsilon_n = C\eta_n^2 \to 0$ for a constant C: it is a counterexample to the floor of Problem 18.2.

Proof. By Lemma 17.7, each model gives a model on $\mathbb{C}^{k_n}$, $1 \le k_n \le D_n$, with $J = -I$ and operator-norm relator errors at most $(1+3L)\eta_n$; with the flat state it is a model of A_G^- at that precision. The upper map of Theorem 5.6 gives a model of the constraint system at precision $C'\eta_n$ on $\mathbb{C}^{12k_n}$ with the flat state, and the Boolean return (Lemma 5.4 and Theorem 4.9(b)) gives a strategy on the maximally entangled state Φ_{12k_n}, the canonical purification of the flat state, with loss at most $C\eta_n^2 =: \varepsilon_n$. Its directional rates are $\log(12k_n) \le \log D_n + \log 12$, so

$$\frac{d_G^{\rightarrow}(\varepsilon_n)}{\log(1/\varepsilon_n)} \;\le\; \frac{\log D_n + \log 12}{2\log(1/\eta_n) - O(1)} \;\longrightarrow\; 0.$$

These strategies show $\omega_q(G) = 1$. A finite-dimensional perfect strategy would give a nonzero finite-dimensional representation of $\mathcal{A}(B_G)$ and, through the unital lower map F, a finite-dimensional representation of the group with $J = -I$, which is excluded. The game is a Boolean constraint-variable game with every incidence tested, so it lies in class (C). □

The proposition is an upper construction on the original resource infimum, through actual pure states; no converse from a small mixed-state E_{D} to a small dimension is used. Its hypothesis also forces $k_n \to \infty$: a bounded dimension along a subsequence would give, by compactness of the unitary groups, an exact finite-dimensional representation with $J = -I$. For a solution group the same holds with the game of the solution group itself and Theorem 4.5 in place of the compiler, by [134, Proposition 4.4(a)]. Appendix F shows which natural candidate sources fail the hypothesis.

18 Open problems

We state the open targets precisely, with what has failed and what would count as progress. The first two are the central ones.

Problem 18.1 (The quarter lower bound)**.** *Prove* $d^{\rightarrow}_{G_{\mathrm{pt}}}(\varepsilon) \geq c\,\varepsilon^{-1/4}$ *eventually, for the exact fixed game of Theorem 13.7, against all original finite-dimensional mixed strategies. By Theorem 8.2 the reverse direction would follow up to a fixed loss rescaling. A bound only on* $S(\rho_B)$*, only on Schmidt rank of pure strategies, or only on the native classes of Section 13.5 is not this result.*

The window is $[\beta, \frac{1}{4}]$ with $\beta = (2\log_2\lambda + 1)^{-1} = 0.0734595360384268649\ldots$. The proof of its lower end balances $N\kappa^2$ against a raw cross-copy bound $\eta_0 \leq C(A+N)e$ from the linear filling charge (Theorem 13.5), so far-commutator areas of order R^u give the exponent $1/(2u+1)$. The geometric gap is open: $R^2/16 \leq A^*(R) \leq CR^p$ for every integer $R \geq 23$, where $p = \log_2\lambda = 6.3064682540119605963\ldots$; the lower bound is [55, Proposition 3.4] and the upper bound is Proposition 13.6a. Quadratic area would give $\varepsilon^{-1/5}$, and cubic area $\varepsilon^{-1/7}$. That is also the limit of this route: the worst-pair raw cross-copy estimate $\eta_0 \leq CN^u e$ is false for every $u < 2$ on the original source models, already at the scale $N \asymp e^{-2/(2u+1)}$ at which the proof uses it, with exact mirrors and only H4 inexact (Theorem G.2 and Proposition G.3). So the raw cross-copy route stops at $u = 2$, that is at $\varepsilon^{-1/5}$, whatever the true growth of the areas, and the quarter needs a different estimate or a different extraction. Proposition 15.4 refutes the analogous uniform raw estimates for lamp commutators below the exponent $\frac{3}{2}$. One different route is a same-state quartic *table* bound $\Delta_N \leq CN^4e^2$, in the regime needed for Proposition 15.5 and with all actual mirrors counted, which would suffice at $N = ce^{-1/2}$ through a different decoder; Proposition 15.6 refutes every smaller power and the raw quartic aggregate shortcut, and the diluted models of Proposition G.3 do not refute the table target, since an averaged table defect sees their bad sector only with its weight. Another original-state operational method would also qualify. An estimate or counterexample that moves the window is what counts; another conditional native layer is not.

Problem 18.2 (The logarithmic floor)**.** *For every fixed game* G *in the classes of Theorem A with value one and no finite-dimensional perfect strategy, are there* $c_G > 0$ *and* $\varepsilon_G > 0$ *with* $d^{\rightarrow}_G(\varepsilon) \geq c_G \log(1/\varepsilon)$ *for* $0 < \varepsilon < \varepsilon_G$*? The constants may depend non-effectively on* G*. A counterexample needs one fixed permitted game and a sequence* $\varepsilon_n \downarrow 0$ *with* $d^{\rightarrow}_G(\varepsilon_n)/\log(1/\varepsilon_n) \to 0$*.*

Every nonattaining game in the permitted classes for which this paper establishes a demand law has at least logarithmic demand; the I_{3322} contrast of Section 16.1 lies outside the permitted format, so it does not refute the floor. The precise state is as follows; every assertion is proved in Section 17 and Appendix F.

A class with the floor. Every game that recovers, with actual mirrors on the original state, a sign relation $WY = -YW$ and an even dilation $XYX^* = Y^m$ has $d^{\rightarrow}_G(\varepsilon) \geq \log(1/\varepsilon)/(4\log m) - O(\log\log(1/\varepsilon))$, with an exponential fixed-target threshold (Theorem 17.3); the Boolean constraint game compiled from Abels' finitely presented solvable group is an instance, with value one and no finite perfect strategy (Corollary 17.6).

A criterion for a counterexample. A finite presentation with a marked central involution, no finite-dimensional representation with $J = -I$, and operator-norm models with $\|U_J - I\|_{\mathrm{op}} > \eta_n \to 0$ and $\log D_n = o(\log(1/\eta_n))$ gives, through the universal compiler, a fixed game of class (C) that violates the floor (Proposition 17.8); the compression of Lemma 17.7 removes any concern about the trace fraction of the negative sector.

Excluded candidate classes. The criterion fails for the following sources. Thin Baumslag–Solitar models, whose operator accuracy is exponential in the dimension: identifying the marker with their commutator kills it (Proposition F.1), and every fixed extension that places the marker in the normal closure of the commutator while retaining copies of the models needs a constant relator error (Proposition F.3 and Example F.4). Locally Schatten-p-stable groups, whose separated models have $\log D \geq p\log(1/\eta) - O(1)$ (Theorem F.5), including the higher-rank Deligne extension of $\mathrm{Sp}(8,\mathbb{Z})$ (Example F.6). Sources containing the dyadic

sign-and-dilation pattern, such as Abels' group, whose separated models have $D \geq c/\eta$ (Proposition 17.5 and Corollary F.7).

Obstructions to the method of Theorem A. The sign holds exactly and the dilation to every accuracy in every type II limit of the proof (Proposition F.8), so their presence in the limit does not distinguish the class of Theorem 17.3, which needs them on the original operators at polynomial cost. Recovering Pauli arrays by bounded polynomials in finitely many tested projections can need a doubly exponential coefficient budget, even with a hyperfinite faithful tracial closure (Proposition F.9). An exponential single-copy extraction threshold suffices for the floor (Proposition F.10), but is not necessary on general grounds: states with distillable entanglement of order $2^{r/2}$ can fail single-copy extraction of r pairs at every fixed infidelity below about $\frac{1}{4}$ (Proposition F.11).

What would reopen or settle the problem. A quantitative original-state extraction or compression theorem that gives the rate for every game of Theorem A, for instance a threshold $\varepsilon_0(G, r, \gamma) \geq \exp(-C_G r)$ at one fixed good γ, would prove the floor; it must produce its decoders from the tested operators with controlled cost, which the type II limit alone does not do. A counterexample must give one fixed game in the permitted classes, of value one and without a finite-dimensional perfect strategy, and original-state upper strategies with sublogarithmic demand along a sequence of losses tending to zero. One sufficient construction is Proposition 17.8. A construction through that proposition must avoid the operator-norm obstructions of Appendix F, including local Schatten stability, the dyadic pattern and copied Baumslag–Solitar models. Those obstructions alone do not rule out a different construction using arbitrary mixed states.

Problem 18.3 (A dense spectrum)**.** *(i) Determine the full set $\mathcal{A}_{\log}$ of usable-entanglement logarithmic coefficients of fixed games of Section 11.6. Every positive Δ_2^0 real is already in this set, on the original mixed state, by [57, Theorems 1.1 and 16.6, Corollary 16.7]; whether the full usable spectrum is larger remains open. (ii) Find fixed games with power laws $d_G^{\rightarrow}(\varepsilon) = \varepsilon^{-\beta+o(1)}$ for a dense set of exponents β in some interval, or with another new asymptotic shape, with complete two-sided operational accounting. (iii) Find a finitely presented exact source with a genuinely new noninteger or intermediate demand law, or a dense interval of laws, with both an original-state lower bound and a finite-rank exact upper construction.*

For (i), the known inclusion is $(\Delta_2^0 \cap (0, \infty)) \subseteq \mathcal{A}_{\log}$, by the matched original-state laws of [57, Theorem 16.6 and Corollary 16.7]. Every element of $\mathcal{A}_{\log}$ is left-c.e. relative to $0''$ (Theorem 11.14); these bounds do not classify the full mixed-state usable spectrum. In contrast, the positive pure original-state entropy spectrum is exactly the positive Δ_2^0 reals (Theorem 11.16). For (ii), finite sums of logarithmic demands remain logarithmic, so $\mathcal{S}$ says nothing about power laws. The only exponents known for the games of this paper are $\beta = 1$, for the configuration-lamp game (Theorem 12.4), and $\beta = 0$, for the logarithmic laws; the point window of Theorem 13.7 does not identify an exponent. For (iii), the subset sources give the positive integers and the configuration and replication sources exponential order. Dimension exponents, logarithmic coefficients and exact-source laws are different quantities, and none of these results settles the others. Proposition 15.9 closes only the finite-support architecture with full stabilizers.

Problem 18.4 (An unrestricted linear compiler)**.** *Construct, for every finite presentation with a marked central involution, a finite linear system game with split algebraic maps and retained-state precision sufficient for the profiles. Proposition 15.3 excludes word decoders, so a polynomial-corner or another architecture is needed; Theorem 5.6 solves the case of Boolean constraint games only.*

Supplying self-contained proofs of the inputs of Theorems 13.11 and 13.12, and of the restricted compiler of Theorem 11.3, would remove their conditions; this is consolidation, not a new discovery. Linear games with every coefficient of Corollary 11.11 already exist by Theorem 11.9, so that compiler matters only for its explicit specifications.

Problem 18.5 (Remaining analytic gaps in the rational family)**.** *(i) Improve the lower remainder of Theorem 11.9 to* $O(1)$*. (ii) Close the* $O(\log\log)$ *gap for* H_∞ *at* $\alpha = 1$ *in Theorem 11.7. (iii) Settle the shared-dilation source with* 4 *and* 9*, presently between rank* $\Omega(e^{-1/2})$ *and* $O(e^{-1})$*; for an exact spectral law invariant under multiplication by* 4 *and* 9*, joint ergodicity and positive entropy for one of these maps imply Haar measure by the theorem of Rudolph and Johnson [125, 76], but that exact-invariance result supplies no stability modulus for the approximately invariant laws of our source models.*

Problem 18.6 (Directional normalization beyond agreement)**.** *Determine the optimal loss and resource conversion between the two directions for games that are not agreement games, or improve the rescaling of demand profiles under a larger, physically specified contract. Theorem 8.3 gives a mechanism boundary with Bob fixed, not optimality for all profile comparisons, and Proposition 8.4 gives only a square-root loss. In a search over small synchronous strategies with non-difunctional off-diagonal predicates no loss ratio above three was found (the largest was* 2.997*), which suggests that the square-root repair is loose.*

Beyond these, two questions inherited from the tool results remain. It remains open whether a state-weighted central-involution profile can be asymptotically much smaller than its flat profile. They agree up to constant precision rescaling for the matching families computed here; the point-source window does not settle this comparison in general. And it is open whether near-optimal strategies for CHSH(n) contain $\Omega(\min(n, 1/\varepsilon))$ distillable ebits; for that game only the entanglement entropy and the entanglement of formation are controlled [101, Theorem E].

19 Discussion

Usable entanglement, not dimension. The question we started from has a cheap version and a costly one. The cheap version asks for a large Hilbert space; a large dimension, or a large marginal entropy, can however sit in a separable state, and, conditionally on [53], the I_{3322} example shows that near a Bell optimum dimension can diverge while usable entanglement stays bounded. The costly version asks whether every near-winning strategy, mixed and correlated with anything, can be made to give up Bell pairs. Theorem A answers it qualitatively for consistency-tested games, and the spectrum answers it quantitatively for the games we can analyse. In both, the conclusion is drawn on the original state (Section 1.6), with actual operators of the two players.

How the proof uses consistency. In the classes of Theorem A every question is checked against a copy of itself or of a coarser question, so near-perfect play forces Alice's operators to mirror Bob's on the state. The mirrors make Bob's moments tracial in the limit and let polynomials found in the limit act back on the actual strategy. For a game in either class with no finite-dimensional perfect strategy, every such limit is of type II: a type I summand would yield a finite-dimensional perfect strategy. A type II algebra contains matrix algebras of every size, and this is where the Bell pairs come from. We do not know that consistency testing is necessary for the conclusion (Section 16).

Three mechanisms. The quantitative results use three mechanisms, each tied to a quantitative invariant of the source. *Counting*: many approximately independent Pauli pairs force entropy and, with extraction, usable entanglement; the pairs come from lamps, and their cost from the areas of the fillings that make lamps commute. *Transport*: dilations and rotations force a spectral law to be close to an invariant one, and invariant laws satisfy a Frostman bound of exponent $\log_m q$, hence have dimension at least $\log_m q$; independent copies add. *Addressing*: under exact relations, the cost is the number of qubits reachable by short words. The gaps between the three are informative. Exact and approximate relations can differ exponentially in cost, because approximate models can merge addresses or replace a sum of defect logarithms by the logarithm of a product (Proposition 13.9); and the quarter exponent of the point game's upper bound sits exactly where collective approximation beats counting.

What is fair to call new. Each result is a composition. The tracial correspondences, finite von Neumann algebra structure, Pauli rigidity and regular dilations, pretty-good recovery, uncertainty with memory and hashing are established. What the compositions add is accounting that keeps the original state at every step, so that the conclusions are lower bounds on the usable entanglement of every mixed near-winner of a fixed game, with explicit rates where we can compute them.

What remains. Two problems are central. The matching lower exponent $\frac{1}{4}$ for the point game would close the only polynomial window we know, $[(2\log_2\lambda+1)^{-1}, \frac{1}{4}]$, with λ from Proposition 13.6a. Bounds by filling area cannot reach it through worst-pair raw cross-copy estimates: that route stops at $\varepsilon^{-1/5}$, the value of a quadratic area bound (Appendix G); it would follow from a same-state quartic bound on a Pauli multiplication table or from another original-state method. The logarithmic floor asks whether every nonattaining consistency game demands at least logarithmic usable entanglement. It holds for every game that recovers a sign and an even dilation with actual mirrors, including a game compiled from Abels' group (Theorem 17.3 and Corollary 17.6). Proposition 17.8 gives an operator-norm route to a counterexample; Appendix F excludes locally stable sources, the dyadic pattern and copied Baumslag–Solitar models from that route. Theorem 17.3 separately excludes games with the stated original-state sign-and-dilation interface. Neither the undecidability of the dichotomy nor the type II structure of the limits decides the problem.

How this work was done

The authors selected the questions and scope, directed the research and take responsibility for every claim. AI models from more than one developer assisted with exploration, derivation, computation and drafting. The models were set against each other in refutation-first checks and repeated whole-paper adversarial reviews, with exact computation. Reports of the checks are available from the authors.

Proof and computational scope. Each result in Table 2 identifies its proof, imported input or conditional hypothesis. The companion-source map in Appendix D records the external results used. The verification scripts test the specified finite constructions and numerical estimates; they do not establish the asymptotic theorems. Appendix C gives reproducibility commands, dependencies and computational limitations. No formal verification of this manuscript is claimed.

Finite computations and exported objects. The scripts in `verify/`, run by `verify/run_all.py`, test the inequalities on random, perturbed, singular and non-flat

instances. They record the qubit count, preliminary lemmas, return theorems, Wasserstein-1 and transport estimates, lamp-source models and orbit arithmetic, random-cut and weighted-history lemmas, cyclic extraction on mixed states with perturbed mirrors, non-orthogonal unrecorded-vector controls, cross-party cut bounds, Fourier lamps with non-dyadic dilations, depth and coefficient arithmetic, balanced lifts and their unbalanced-marker controls, the universal compiler's exact split identity, finite linear games and tree-of-faces identities. The source relators and their raw finite words are tested before group labels are substituted. These are checks of the enumerated instances, not general proofs.

The larger exported objects are archived in `verify/artifacts/` with SHA-256 hashes. The runner `verify/run_artifacts.py` replays their checks: the point-Pauli Boolean game against all 22 source relators; the archived L4M8 game against its presentation, squares, rows and row commutators using the normal-form engine of `check_linear_games.py`, with complete Fourier-lamp play at four sizes; the two-copy game in the negative source algebra, with every context and raw outcome played at two sizes; the selector construction on two small inputs using L4M8 in place of the fallback game; and bounded exact controls for the literal subset and replication sources. The archived-game checker shares that normal-form engine, so it is not an independent implementation of the engine.

Spectrum, floor and window checks. The script `verify/check_refinements.py` checks nineteen finite groups of assertions, including arithmetic arrays, flagged ensembles, physical leakage, floor diagnostics, Laurent freeness and the small-error pairwise budget. The current point-window exponent is recorded by `check_tree_filling.py` and the window engines. The script `verify/check_window.py` runs two independently implemented free-word verifiers of one certificate set produced by the first engine. Each verifies thirty-seven core and six transfer certificates before and after literal auxiliary-relator expansion; the second also checks sixty holonomy examples. They are two verifiers of the same certificate set, not two independent certificate-generation procedures. Their recorded JSON contains the matrices, spectra, constants and certificate results. The runners execute one process at a time with one numerical-library thread. Rerunning intentionally refreshes the JSON and logs specified in the repository README.

The leakage proof uses Winter's conditional-entropy bound [147, Lemma 2]. The magnetic-cycle spectrum and equal-distribution minimization used for holonomy are credited to [40, Lemma 2.9] and [21, Proposition 4.6(b)]. Recovery, exact synchrony, hashing, convex roofs and entropy continuity retain their cited sources at use.

Literature and limits. Bounded primary-literature searches compare the headline statements with the works credited at the results; they do not establish exhaustive priority. The white-box protocol is compared with universal distillation [135, 138], compound-source distillation [20] and relation decomposition [99]. Comparisons also include self-testing [8, 143, 31], the I_{3322} literature [119, 37], and the tools listed in Section 1.7. Written proofs remain the evidence for the asymptotic statements; finite computations and reviews support their checking.

A Proof of the white-box protocol

We prove Theorem 7.7 for class (S); the changes for class (C) are at the end.

The interface. Each party holds the physical payload of every copy and may append known pure registers. For each question x there is a reusable unitary V_x, with its inverse, acting on the payload and a retained local Naimark register, and known outcome projectors Q_a

define the dilated PVM $P_{xa} = V_x^* Q_a V_x$, which on initialized Naimark registers reproduces the strategy's statistics. A controlled reflection $2P_{xa} - I$ is implemented as V_x, then a controlled known phase $2Q_a - I$, then V_x^*; no controlled version of the unknown V_x is needed. Naimark registers are retained across word queries, never reset. The same state and the same unitaries are used on every copy.

Thresholds. Let $F^{(1)}, \dots, F^{(K)}$ be the menu of Theorem 7.5 at threshold b, so that every strategy of loss at most $\varepsilon_{\mathrm{cert}}$ has an array with $D_B(F) = \langle \psi, d_F(P^B)\psi \rangle < b$. Let $K_* \geq 1$ bound $\sum_w |c_w||w|$ for every F_j and every d_F, and put $\varepsilon_* = \min\{\varepsilon_{\mathrm{cert}}, (\mu_{\min}/2)(b/K_*)^2\}$. Then every raw polynomial mirror, including those of the d_F, has error at most b by (6.1), so $|D_A(F) - D_B(F)| \leq b$, where $D_A(F)$ is the expectation of $\theta(d_F)$. Since the pair defect $C = F_j F_k - s F_k F_j$ satisfies $C^* = -sC$, it is normal and $\theta(C^*C) = C_A^* C_A$; so $\theta(d_F)$ is the positive sum of Alice's own square and pair defects.

Alice-only selection. Expand each d_F, after $P = \frac{1}{2}(I + R)$, as a rational combination of words in the reflections R_{xa}, splitting complex coefficients into real and imaginary parts with phases absorbed into the unitary terms; let Λ_D be a common bound on the ℓ^1 norms of the coefficients. On a sacrificed copy Alice picks a term with probability proportional to its coefficient, runs its controlled-unitary Hadamard test and records a sign whose expectation is $D_A(F)/\Lambda_D$. With $N = \lceil 2\Lambda_D^2 b^{-2} \ln(2K/\zeta) \rceil$ copies per array, Hoeffding's inequality and a union bound give, with probability at least $1 - \zeta$, all K estimates within b. Alice announces the index of a smallest estimate. A Bob-good array has $D_A < 2b$ and estimate below $3b$, so the selected array has $D_A \leq 4b$ and $D_B \leq 5b$: square and pair residuals at most $3\sqrt{b}$ on both parties, raw mirrors at most b. Because the copies are independent and the oracles memoryless, conditioning on the selection transcript leaves the retained copy's AB marginal unchanged; its global purification may be steered by measurements on other copies, which is harmless because all bounds below hold for every purification and the environment and transcript are kept as junk.

Coherent sign rounding. For one F with coefficient bound Λ, a known PREPARE and a term-by-term SELECT give a unitary B with top block F/Λ; its oracle cost is the sum L_F of the term lengths. With $V = |0\rangle\langle 1| \otimes B + |1\rangle\langle 0| \otimes B^*$, signal vector $s = |+\rangle \otimes |0\rangle$, $P = |s\rangle\langle s| \otimes I$ and $W = (2P - I)V$, the walk W has, on the invariant plane of an eigenvector of F with eigenvalue λ, eigenvalues $e^{\pm i\vartheta}$ with $\cos\vartheta = \lambda/\Lambda$, so the cosine of either phase has the sign of λ when $\lambda \neq 0$. Phase estimation with $Q \geq 1024\Lambda$ points, classified by the sign of the cosine and repeated an odd number $J \geq 4\log(4/b)$ of times with majority vote, has classification error at most $b/4$ outside the window $|\lambda| < \frac{1}{2}$; applying the majority sign as a phase and uncomputing gives an exact reflection T on payload and work registers with $\|(T - \mathrm{sgn}(F) \otimes I)(\psi \otimes |0\rangle)\| \leq \sqrt{b} + \frac{8}{3}\|(I - F^2)\psi\| \leq 9\sqrt{b}$. The second term bounds the weight of the excluded window. Each T_j has its own work register, retained when dirty; T_j is an exact reflection on the whole extended space, and controlled T needs only control of the known classifier phase.

Dirty work does not spoil the decoder. The ideal reflections $Q_j = \mathrm{sgn}(F_j)$ satisfy, by product replacement with opposite-party mirrors as in Step 4 of Theorem 7.2, every Pauli pair and mirror bound with constant $20\sqrt{b}$. Replacing each Q_j by the implemented T_j inside a product costs at most $9\sqrt{b} + 40\sqrt{b}$ per factor, again by moving the right factor to the opposite party; so the implemented reflections satisfy all pair and mirror bounds with $\kappa = 200\sqrt{b}$. Both parties then apply the signed regular isometry of Lemma 6.5 to the T_j, using m controlled T calls. Its Bell-stabilizer errors are at most $(2C_r + 1)^2\kappa^2/4$ and the singlet

infidelity at most $2r(2C_r+1)^2\kappa^2/4 = 20000\,r(2C_r+1)^2 b < \gamma/1000$. Adding the selection failure $\zeta = \gamma/8$ keeps the infidelity below $\gamma/4$, and finite-precision compilation of all known gates, with per-gate accuracy below $\gamma/[100(T_{\mathrm{all}}+1)]$ for the total count T_{all} of invocations, changes the output by less than $\gamma/10$ in trace distance. The only cross-lab message is the menu index, of $\lceil \log K \rceil$ bits, from Alice to Bob.

Class (C). Use Theorem 7.5 for native binary constraint games, with the thresholds there. For each variable j fix an incident context $i(j)$; Alice's controlled proxy reflection is $V_{i(j)}$, a known phase on the labels whose j-th bit has the queried value, and $V^*_{i(j)}$. Illegal padded labels are relabelled virtually to a fixed legal assignment by known phases on the outcome register. Distinct proxies need not commute, and nothing assumes they do; the rest of the proof is unchanged, with $\mu_{\min}$ the least incidence weight.

B Status of the results

Table 2 lists every result of the paper with its status. *Proved*: a complete proof is given here, using only the cited results named in the row or the proof. *Checked*: a finite computation or a verified implication from a cited manuscript. *Conditional*: the result rests on a named input not proved self-containedly. *Companion*: the result is stated here and proved in [101]. The final column describes the proof or finite computational scope. The references in the location column identify the written proofs; Appendix C gives commands, arithmetic conventions and limitations for the computations. Finite examples do not substitute for the proofs.

result	where	status	proof / finite scope
qubit count in a state	Theorem 3.1 and Corollary 3.2	companion	Cited companion proof
transfer and return, linear	Theorems 4.3 and 4.5 and Corollary 4.6	companion	Cited companion proof
Boolean correspondence	Theorem 4.9	proved	Written proof
universal Boolean compiler	Theorem 5.6	proved	Exact split identity
compiler on virtual retracts	Theorem 5.8	proved, on [101, Lemma B.1]	Written proof
balanced compiler, extended presentations	Proposition 5.10	proved	Finite lift controls
actual source interface	Proposition 6.1	proved	Written proof
triangular extraction	Theorem 6.2	proved	Written proof
cyclic triangular extraction; selector polynomials	Lemma 6.3, paragraph after Proposition 6.1	proved	Mixed-state tests
signed regular extraction	Theorem 6.4 and Lemma 6.5	proved	Written proof
aggregate and finite-group extraction	Theorems 6.6 and 6.7, Corollary 6.8, and Lemma 6.9	proved	Written proof
Alice-local recovery, forward bound	Theorem 6.10	proved	Written proof
exact synchrony	Theorem 6.11	proved	Written proof

result	where	status	proof / finite scope
optimal small-error pairwise budget	Theorem 6.12	proved	Finite budget tests
Theorem A	Theorem 7.2	proved	Written proof
effective certificates	Theorem 7.5	proved	Written proof
undecidability	Theorem 7.6	proved	Finite selector tests
white-box protocol	Theorem 7.7, Appendix A	proved	Written proof
local symmetry, Bob kept	Theorem 8.1	proved	Written proof
factor three and sharpness	Theorems 8.2 and 8.3	proved	Written proof
synchronous repair; directions	Proposition 8.4 and Corollary 8.5	proved	Written proof
Slofstra–Vidick dimension and profile	Theorems 9.3 and 9.9 and Corollary 9.4	companion; upper lift [101, Lemma 11.13]	Finite examples
Slofstra–Vidick entropy and formation	Theorems 9.6 and 9.10 and Corollary 9.8	proved	Written proof
one lamp source	Theorem 10.2	proved	Written proof
additive profiles	Theorem 10.9	proved	Written proof
shared dilations $4, 6$	Theorem 10.13	proved	Written proof
rational Boolean constraint games	Theorem 11.1	proved	Written proof
one-copy linear family	Theorem 11.2	proved	Written proof
restricted compiler, several copies	Theorem 11.3	conditional	Named hypotheses
linear games, every rational exponent	Remark 11.5 and Corollary 11.6	proved	Written proof
Rényi spectrum, concentration	Theorem 11.7	proved	Written proof
sharp E_{D} coefficient, rational α	Theorem 11.9	proved	Written proof
sharp E_{D} coefficient, every $\alpha \in \mathcal{S}$	Theorem 11.9 and Remark 11.10	proved	Written proof
dense coefficient set; arithmetic bounds	Corollary 11.11, Lemma 11.12, Proposition 11.13, and Theorem 11.14	proved	Written proof
pure coefficient spectrum	Theorem 11.16	upper class proved here; lower inclusion [57, Theorem 1.1]	Written proof; cited realization
flagged coefficient spectrum; matched extraction	Propositions 11.15 and 11.17	proved; extraction interface [57, Lemma 12.1]	Written proofs; finite spectrum checks

result	where	status	proof / finite scope
mixing, computable fixed states and abstract arithmetic	Propositions 11.18 to 11.20	proved	Finite spectrum checks
weak rounding from matched laws	Proposition 11.21	proved; realizations [57, Theorem 1.1 and Section 12]	Written proof; cited matched laws
formation, synchrony and physical leakage	Propositions 11.22 to 11.24	proved	Finite spectrum checks
kernel, half-erasure and nonlimit obstructions	Propositions 11.25 to 11.27	proved	Finite spectrum checks
configuration-lamp game: entropy, formation, dimension	Theorems 12.1 and 12.3	companion	Cited companion proof
configuration-lamp game: usable entanglement	Theorem 12.4	proved	Written proof
linear filling charge	Theorem 13.5	proved	Tree-of-faces tests
point window; six-type fillings	Proposition 13.6a, Theorem 13.7 and Proposition 13.9	proved	Exact word certificates
fixed-holonomy subdivision	Proposition 13.9a	proved from magnetic-cycle estimate	Cycle examples
native class laws	Theorems 13.11 and 13.12	conditional	Named hypotheses
address growth	Theorem 14.1	proved	Written proof
exact point source and traces	Theorem 14.3 and Propositions 14.2 and 14.4	proved	Written proof
exact configuration source	Theorem 14.5	proved	Written proof
subset sources	Proposition 14.6 and Theorems 14.7 and 14.8	proved, on Johnson's presentation	Written proof
replication source	Theorem 14.9	proved, on Johnson's presentation	Written proof
PPT gap; no exact output	Theorem 15.1 and Remark 15.2	proved	Written proof
linear compiler exclusions	Proposition 15.3	proved	Written proof
raw pair bounds	Proposition 15.4	proved, on the regions of [55, §6]	Written proof
finite-table gate; refutations	Propositions 15.5 and 15.6	proved	Written proof
correction costs	Proposition 15.7	proved; native part conditional	Named hypotheses
scalar phase schemes	Proposition 15.8	proved	Written proof

result	where	status	proof / finite scope
finite supports; invisible witnesses	Propositions 15.9 and 15.10	proved	Written proof
I_{3322} Bell-gap contrast	Theorem 16.1	checked; conditional on [53]	Named hypotheses
floor on the sign-and-dilation class; Abels game	Theorem 17.3 and Corollary 17.6	proved	Written proof
compression; dyadic bound; criterion	Lemma 17.7 and Propositions 17.5 and 17.8	proved	Written proof
excluded candidate classes	Propositions F.1 and F.3, Theorem F.5, and Corollary F.7	proved	Written proof
Deligne extension	Example F.6	proved from imported results [50, 5, 47]	Written proof
method obstructions	Propositions F.8 to F.11	proved	Written proof
continuous amplification; spectral multiplicity; polynomial augmentation	Propositions F.12 to F.14	proved	Finite refinement checks
source leakage, stability and populated slow modes	Propositions F.15 to F.17	proved	Finite refinement checks
qualitative thresholds; selectable depth	Remark F.18	proved	Finite refinement checks
raw cross-copy obstruction	Theorem G.2 and Proposition G.3	proved	Written proof
finite blocks	Remark 16.2	known [83]	Cited result

Table 2: Status of every result.

C Reproducibility and computational scope

The mathematical evidence is the written proof at each result and its explicitly cited external inputs. Table 2 records proof locations and the associated finite computational scope; Appendix D identifies the companion results. The programs below check specified finite instances and exported objects. They do not prove an asymptotic estimate, decide an unrestricted optimization problem, or constitute a formal verification of this manuscript.

Dependencies and execution. The checks require Python 3, NumPy and SciPy. Run the following commands from the repository root, one runner at a time. Set each of these environment variables to `1`:

```
OMP_NUM_THREADS, OPENBLAS_NUM_THREADS, MKL_NUM_THREADS
```

```
NUMEXPR_NUM_THREADS, VECLIB_MAXIMUM_THREADS
BLIS_NUM_THREADS, PYTHONDONTWRITEBYTECODE
```

The README gives the corresponding shell commands.

```
python -I -X utf8 -B verify/run_all.py
python -I -X utf8 -B verify/run_artifacts.py
```

The first runner writes one log per check under `verify/logs/`, with structured outputs where specified by each script. The second records its steps under `verify/logs/artifacts/` and writes the exported-object results beside the corresponding scripts. Rerunning intentionally refreshes those outputs. The input manifest `verify/artifacts/SHA256SUMS` is checked before replay; refresh it only when an input is intentionally changed.

Exact computations. The artifact checks use integer and rational arithmetic, literal free words, finite permutations and the stated normal-form procedures. They compare the point-game source relators with those printed in the paper, and check the exported variables, contexts, squares, products, commutators and recovery maps. The archived L4M8 checker tests the actual game file using the normal-form engine of `check_linear_games.py`; those programs share that engine. The two-copy checker works in the negative source algebra. The selector checker verifies its finite conditional-parity predicates, incidence weights and an explicit perfect strategy. Subset and replication checks test the literal source presentations on bounded exact instances. These are finite identities and constructions; their extension to arbitrary strategies uses the written proofs.

Floating-point computations. The numerical suites test matrix identities, norms, entropies, transport inequalities, mixed-state extraction and explicit lamp models at their recorded dimensions and tolerances. They include singular and non-flat states, perturbed mirrors and negative controls. The refinement suite checks arithmetic arrays, flagged ensembles, physical leakage, floor diagnostics, Laurent freeness and the small-error pairwise budget on the listed finite instances. Numerical eigenvalues, entropy values and fitted slopes are floating-point observations, not interval certificates or universal error bounds. Where high-precision arithmetic is used, the recorded precision is part of the computational scope.

Word certificates and the point window. The two window engines independently verify one common set of word certificates, produced by the first engine. Each checks the core and transfer identities before and after literal auxiliary-relator expansion; the second also checks finite holonomy examples. They are separate verifiers of the same certificates, and do not supply two independent generation procedures. The recorded matrices, spectra, constants and certificate outcomes support reproducibility. The filling and original-state extraction arguments in the paper supply the asymptotic conclusion.

Building the paper. With plain pdflatex and BibTeX installed, run:

```
python paper/build.py
```

The helper performs four pdflatex passes with BibTeX between the first and second, disables shell escape, and keeps intermediate files outside the release tree. It requires a final log without warnings, undefined references or overflowing boxes. The arXiv folder contains the active TeX inputs and the matching BBL, and builds with plain pdflatex without the verification programs.

Limits of the evidence. Finite tests do not establish compactness arguments, convergence rates, arbitrary-state lower bounds, the full group normal-form theorems, the universal game reduction or the operational entanglement-rate theorems. A negative control demonstrates detection of the stated alteration, not all possible defects. Imported theorems and conditional hypotheses remain identified at their uses. No formal verification of this manuscript is claimed. Reports of the checks are available from the authors.

D Results used from our companion papers

Table 3 lists every result of our companion papers [101, 54, 55, 56, 53, 57] on which an argument here depends, and where it is used. Results mentioned only for comparison are marked as such. The results of [101] listed here are stated in this paper with a pointer to their complete proofs there. The table identifies the precise imported results used by each argument.

result	source	used in
compressing a strategy to the Schmidt supports	[101, Rem. 2.9]	Section 2.1
standard form, consistency, calculus in a state, conjugator lengths, rounding, rotating the state	[101, Lemmas 3.1–3.4, 3.6, 3.7]	Section 2
projection products; two-sided twirls; sector states; entropy facts	[101, Lemmas 3.8–3.10, §3.7], [55, Lemma 4.1]	Section 2
qubit count in a state; uniform errors; Jensen–Shannon bound; entropy of a sector	[101, Thm. 4.1, Cor. 4.2, Lemmas 4.8, 4.9]	Section 3
CHSH(n) pure entropy and formation (comparison)	[101, Thm. E]	Remark 3.3
transfer, return and two-sided correspondence for linear system games	[101, Thms. 5.1, 5.4, Cor. 5.5]	Section 4
measures on the circle; coupling in one state	[101, Lemmas 11.7, 11.8]	Section 9
profile of c in a state; spectral measure of y in a state; flat model; profile of c; balanced lift for K; the Slofstra–Vidick game	[101, Thm. 11.9, Prop. 11.10, Ex. 11.11, Cor. 11.12, Lemma 11.13, Thm. 11.14]	Section 9
$E\Gamma_0$ presentations	[101, Def. 10.1]	Theorem 5.8
the four gadget steps; Ψ and x_{j_0} with (G1), (G2); block words	[101, Lemma B.1, Table 3]	Theorems 5.6 and 5.8, Section 12
embedding theorem, flat and in a state (comparison)	[101, Thm. B.3]	after Theorem 5.8
Houghton's group, its action, its map to $\mathbb{Z}^2$ and the presentation P_H	[101, Def. 6.1, (5), (6), App. A]	Section 14.4, Theorem 14.5
the certificate R_{P}	[101, Def. 7.1, Table 2]	Section 12, Theorem 14.5
point stabilizers; transporters	[101, Lemma 7.4, Def. 7.5]	Theorem 14.5
pair areas and their growth	[101, Thm. 7.8, Rem. 7.10]	Section 12, Theorem 14.5
transporting J	[101, Lemma 7.11]	Section 12, Theorem 14.5

result	source	used in
the sector theorem, its proof's estimates, and its polynomial form	[101, Thm. 8.1, Rem. 8.3]	Section 12
amenability of G_{P}, the quotient $\mathbb{Z}^2$	[101, Prop. 6.8]	Theorem 14.5
the clock lemma	[101, Lemma 9.3]	Theorem 14.5
hash models (comparison)	[101, Prop. 9.4]	Section 1.3, Proposition 15.7
the presentation $\mathcal{P}$, words $\hat{s}$, area $A_* = 60$	[101, Prop. 10.3]	Section 12
the game G, $m = 1687$, $s = 3$; entropy, formation and upper bounds	[101, Thm. A, Thm. 10.5, Cor. 10.6]	Section 12
the S_3 configuration lamps: profile, unitary models, weighted one-round theorem (comparison)	[54, Thm. A, Thm. 10.1]	Sections 1.7 and 12
actual-state extraction	[57, Lemma 12.1]	Proposition 11.15
matched laws and weak rounding	[57, Section 12]	Proposition 11.21
limit-computable realization	[57, Theorem 1.1]	Problem 18.3
matched logarithmic dimension and original-state usable laws	[57, Thm. 16.6, Cor. 16.7]	Section 11.6 and Problem 18.3
Johnson's relators in a, b, α	[55, (1)]	Section 13.1
far-commutator filling method	[55, Thm. 3.2, App. A]	Proposition 13.6a and Appendix G.1
retained-relator refinement framework	[55, Rem. A.3]	Appendix G.1
quadratic lower bound for far commutators (comparison)	[55, Prop. 3.4]	Section 18
polynomial Dehn function (comparison)	[55, Thm. 3.3]	Remark 12.2
corner models; rotating links; regions and spanning trees; spreading a defect	[55, Prop. 6.1, Lemmas 6.3, 6.4, proof of Prop. 6.5]	Proposition 13.9, Appendix G
unitary models, $14.26M^{-3}$	[55, Prop. 6.5]	Proposition 13.9, Section 13.5
sharp Houghton cost with moderate flavour growth (conditional input)	[56, App. R]	Theorems 13.11 and 13.12
nonattainment, spatial attainment, rate, geometric tails	[53, Thms. 1, 2, 6, §5.2]	Theorem 16.1

Table 3: Results of our companion papers used here. Numbers refer to version 1.0.0 of [101], arXiv version 1 of [54], version 1.1.2 of [55], version 1.0.0 of [56], version 3.0.0 of [57] and arXiv version 2 of [53], whose content is that of its version 4.1.0.

E The Ji–Leung–Vidick game with a tapered reservoir

Ji, Leung and Vidick built a three-player game, 3EMB, in which near-optimal play forces the players to exchange an EPR pair coherently with a product state, which needs an embezzling reservoir [74]; embezzling states go back to van Dam and Hayden [141]. Ji, Leung and Vidick proved that winning with probability $1-\varepsilon$ needs local dimension $2^{\Omega(\varepsilon^{-c})}$ for two of the players, and that $O(\varepsilon^{-1})$ qubits per player suffice [74, Theorem 1]. Their strategy uses the reservoir $|\Gamma_d\rangle = N_d^{-1}\sum_{j=1}^{d}|11\rangle^{\otimes j}|\mathrm{EPR}\rangle^{\otimes(d-j)}$, whose left-shift unitaries perform the exchange with overlap $1-O(1/d)$ [74, (7)]. A reservoir whose amplitudes taper to zero at both ends does the exchange with overlap $1-O(1/d^2)$. Such amplitudes, proportional to $\sin(\pi k/(N+1))$, appear in the coherent state exchange of Leung, Toner and Watrous, who credit the improvement to Harrow and Shor [85, footnote 1]; see also [86]. This appendix records that substituting a

tapered reservoir into the strategy of [74, Lemma 7] gives the upper bound $O(\varepsilon^{-1/2})$ qubits. We claim no new reservoir and no matching lower bound.

Lemma E.1 (A tapered reservoir)**.** *Let $L \geq 1$ and $P_m = \frac{2}{L+1}\sin^2\frac{\pi(m+1)}{L+1}$ for $0 \leq m < L$. Let $\Gamma_L \in (\mathbb{C}^2)^{\otimes L} \otimes (\mathbb{C}^2)^{\otimes L}$ have Schmidt coefficients whose squares are $P_m/2^m$, each with multiplicity 2^m, for $0 \leq m < L$. There are unitaries U and V on $\mathbb{C}^2 \otimes (\mathbb{C}^2)^{\otimes L}$ such that*

$$F = \langle |11\rangle \otimes \Gamma_L,\ (U \otimes V)(|\mathrm{EPR}\rangle \otimes \Gamma_L)\rangle$$

is real and at least $\cos\frac{\pi}{L+1}$.

Proof. The P_m sum to 1, since $\sum_{k=1}^{L}\sin^2(\pi k/(L+1)) = (L+1)/2$, and Γ_L has Schmidt rank $2^L - 1$, so it fits in L qubits on each side. The state $|\mathrm{EPR}\rangle \otimes \Gamma_L$ has, for each m, 2^{m+1} Schmidt coefficients with squares $P_m/2^{m+1}$, and $|11\rangle \otimes \Gamma_L$ has 2^{m+1} Schmidt coefficients with squares $P_{m+1}/2^{m+1}$ in shell $m+1$. Let U map the Schmidt vectors of the first state on Alice's side in group m to those of the second in shell $m+1$, for $0 \leq m \leq L-2$, and complete it to a unitary by an arbitrary bijection of the remaining vectors of orthonormal bases; let V do the same on Bob's side with the same bijection. Then F is a sum of products of nonnegative Schmidt coefficients, so it is real and nonnegative, and the matched groups alone contribute

$$\sum_{m=0}^{L-2} 2^{m+1}\sqrt{\frac{P_m}{2^{m+1}}\cdot\frac{P_{m+1}}{2^{m+1}}} = \frac{2}{L+1}\sum_{j=1}^{L}\sin(j\vartheta)\sin((j+1)\vartheta) = \cos\vartheta,$$

with $\vartheta = \pi/(L+1)$; the term $j = L$ vanishes, and $\sum_{j=1}^{L}\sin(j\vartheta)\sin((j+1)\vartheta) = \frac{1}{2}(L\cos\vartheta - \sum_{j=1}^{L}\cos((2j+1)\vartheta)) = \frac{L+1}{2}\cos\vartheta$. □

Proposition E.2. *For every $L \geq 1$ the game* 3EMB *of [74, Figure 2] has a strategy in which the first player holds three qubits and the other two hold $L+3$ qubits each, and which wins with probability at least $1 - \pi^2/(16(L+1)^2)$. Hence $O(\varepsilon^{-1/2})$ qubits per player suffice to win with probability $1-\varepsilon$.*

Proof. Take the strategy of the proof of [74, Lemma 7] with Γ_L in place of Γ_d, and with the controlled unitaries U and V of Lemma E.1 in place of the controlled left shifts. In that strategy the reservoir is used only in part (d); in parts (a), (b) and (c) the second and third players act on their GHZ registers only, the same way in all three parts, and the strategy is accepted with probability 1 [74, proof of Lemma 7]. Part (d) is played with probability $\frac{1}{4}$, and there the referee always accepts if the first player reports $u = 1$, which happens with probability $\frac{1}{2}$. Conditioned on $u = 0$, the state is [74, (10)]

$$\tfrac{1}{\sqrt{2}}|0\rangle_{V_1}\left(|0\rangle_{V_2}|00\rangle_{A_1B_1}|00\rangle_{A_2B_2} + |1\rangle_{V_2}|\mathrm{EPR}\rangle_{A_1B_1}|11\rangle_{A_2B_2}\right) \otimes |\mathrm{GHZ}\rangle \otimes \Gamma_L.$$

After the second and third players apply U and V to A_1A' and B_1B', controlled by A_2 and B_2, the inner product with the ideal state, in which $|\mathrm{EPR}\rangle_{A_1B_1}$ is replaced by $|11\rangle_{A_1B_1}$, is $\frac{1}{2}(1+F)$, a positive real number. The remaining operations of the strategy are unitaries applied to both states, followed by the honest measurements of the test, which accept the ideal state, the state [74, (12)], with certainty. If an effect $0 \leq E \leq I$ satisfies $\langle v, Ev\rangle = 1$ for a unit vector v, then $Ev = v$ and $E \geq |v\rangle\langle v|$. So, conditioned on $u = 0$, the strategy is rejected with probability at most $1 - \frac{1}{4}(1+F)^2 \leq 1 - F \leq \frac{\pi^2}{2(L+1)^2}$, and the total failure probability is at most $\frac{1}{4}\cdot\frac{1}{2}\cdot\frac{\pi^2}{2(L+1)^2}$. Taking $L = \lceil \pi/(4\sqrt{\varepsilon})\rceil$ gives the last claim. □

So the number of qubits that the players of 3EMB need lies between $\Omega(\varepsilon^{-c})$ and $O(\varepsilon^{-1/2})$. A single coherent exchange cannot do better than Lemma E.1: by [74, Fact 8], which is [122, Fact 5.7], an exchange with infidelity η needs a reservoir of local dimension t with $\log t = \Omega(\eta^{-1/2})$. A matching lower bound for the game would need a soundness argument in which the exchange infidelity is linear in the losing probability; we do not have one.

F The floor: candidate classes and method obstructions

This appendix proves the statements about the logarithmic floor that Problem 18.2 summarizes. The first four subsections concern candidate sources. The remaining subsections study method obstructions, continuous amplification, polynomial augmentation and original-state kernel leakage. Throughout, operator norms are written $\|\cdot\|_{\mathrm{op}}$, and $\|\cdot\|_{2,k}$ is the normalized Hilbert–Schmidt norm on $\mathbb{C}^k$.

F.1 Centralizing a Baumslag–Solitar commutator kills it

The Baumslag–Solitar relation $vu^2v^{-1} = u^3$ has exponentially accurate unitary models of small dimension, whose commutator witness $W = [u, vuv^{-1}]$ is visible in operator norm (Example F.4). Making that witness itself the central marker kills it.

Proposition F.1. *Let A, B be unitaries, $W = ABA^*B^*$, $\delta = \|B^2 - A^3\|_{\mathrm{op}}$, $\gamma_A = \|WA - AW\|_{\mathrm{op}}$ and $\gamma_B = \|WB - BW\|_{\mathrm{op}}$. Then $\|W - I\|_{\mathrm{op}} \le 4\delta + 3\gamma_A + \gamma_B$. If $B = VAV^*$ for a unitary V with $\gamma_V = \|WV - VW\|_{\mathrm{op}}$, then $\|W - I\|_{\mathrm{op}} \le 4\delta + 4\gamma_A + 2\gamma_V$. Consequently, if a unitary J has $\|J - W\|_{\mathrm{op}}$, $\|JA - AJ\|_{\mathrm{op}}$, $\|JV - VJ\|_{\mathrm{op}}$ and δ all at most τ, then $\|J - I\|_{\mathrm{op}} \le 23\tau$.*

Proof. The identities

$$AB^2A^*B^{-2} = W\,(BWB^*), \qquad A^3BA^{-3}B^{-1} = (A^2WA^{-2})(AWA^{-1})W$$

hold in every group. In the first, replacing B^2 by A^3 twice turns the left side into I at cost 2δ, and replacing BWB^* by W costs γ_B; so $\|W^2 - I\|_{\mathrm{op}} \le 2\delta + \gamma_B$. In the second, replacing A^3 by B^2 twice costs 2δ, and the conjugates of W by A^2 and A differ from W by at most $2\gamma_A$ and γ_A; so $\|W^3 - I\|_{\mathrm{op}} \le 2\delta + 3\gamma_A$. Since $W = W^3(W^2)^*$, $\|W - I\|_{\mathrm{op}} \le \|W^3 - I\|_{\mathrm{op}} + \|W^2 - I\|_{\mathrm{op}}$. If $B = VAV^*$, then $WB - BW = (WV - VW)AV^* + V(WA - AW)V^* + VA(WV^* - V^*W)$, so $\gamma_B \le \gamma_A + 2\gamma_V$. For the last claim, $\gamma_A \le \|JA - AJ\|_{\mathrm{op}} + 2\|J - W\|_{\mathrm{op}} \le 3\tau$ and likewise $\gamma_V \le 3\tau$, so $\|W - I\|_{\mathrm{op}} \le 4\tau + 12\tau + 6\tau$ and $\|J - I\|_{\mathrm{op}} \le 23\tau$. □

So no extension that identifies the central marker with the commutator, up to a vanishing error, keeps the marker separated from I; the relations $W^2 = 1$ and $W^3 = 1$ that centrality forces are coprime.

F.2 Clocked witnesses and copied normal-closure gadgets

A subtler extension keeps the old Baumslag–Solitar matrices and puts the marker into the normal closure of their commutator. The following two statements show that this fails at constant error whenever the old witness lives on one cell of a clock.

Lemma F.2 (Clock projections)**.** *Let a unitary V cyclically permute orthogonal projections $E_0, \dots, E_{n-1}$ with $\sum_j E_j = I$, that is $VE_jV^* = E_{j-1}$ with indices modulo n, of arbitrary ranks. Let P be a nonzero projection of finite rank k, $p_j = \mathrm{Tr}(PE_j)/k$ and $\gamma = \|PV - VP\|_{\mathrm{op}}$. Then $\max_j p_j \le 1/n + \gamma$.*

Proof. Measuring $\{E_j\}$ in the state VPV^*/k gives the probabilities $\mathrm{Tr}(VPV^*E_j)/k = \mathrm{Tr}(PE_{j+1})/k = p_{j+1}$, a cyclic shift of (p_j). Measurement contracts the trace norm, so $\sum_j |p_{j+1} - p_j| \le \|P - VPV^*\|_1/k \le 2\|P - VPV^*\|_{\mathrm{op}} = 2\gamma$, using $\mathrm{rank}(P - VPV^*) \le 2k$. A cyclic sequence has total variation at least twice its range, so $\max_j p_j - \min_j p_j \le \gamma$, and $\min_j p_j \le 1/n$. □

Proposition F.3 (Copied normal-closure gadgets)**.** *Let U, V be unitaries on $\mathbb{C}^d$ such that V cyclically permutes orthogonal projections $E_0, \dots, E_{n-1}$ summing to I, and $W = [U, VUV^*]$ satisfies $(W - I)^*(W - I) \le 4E_0$. Let $\mathcal{P} = \langle S \mid R \rangle$ be a finite presentation containing generators u, v, J and the relators J^2 and $[J, s]$ for $s \in S$, with largest relator length L, and suppose that J is freely equal to a product of $a \ge 1$ conjugates of $w^{\pm 1}$, $w = [u, vuv^{-1}]$, and b conjugates of relators. Every unitary assignment with $u \mapsto U \otimes I_q$, $v \mapsto V \otimes I_q$ for some $q \ge 1$, operator-norm relator errors at most τ and $\|U_J - I\|_{\mathrm{op}} > \tau$ satisfies*

$$2 \;\le\; 2a\sqrt{1/n + 3\tau/2} + (24a + b(1 + 3L))\tau.$$

In particular, if $n \ge 8a^2$, then $\tau \ge \min\{1/(24a + b(1 + 3L)),\, 1/(12a^2)\}$.

Proof. Lemma 17.7 gives unitaries Q_s with $\|Q_s - U_s\|_{\mathrm{op}} \le 3\tau$, all commuting with a nonzero projection P of rank k, with $Q_J = -I$ on the range of P and relator errors at most $C\tau$, $C = 1 + 3L$. Its proof also gives $\|[P, U_v]\|_{\mathrm{op}} = \frac{1}{2}\|[H, U_v]\|_{\mathrm{op}} \le \frac{3}{2}\tau$. Apply Lemma F.2 to $V \otimes I_q$ and the projections $E_j \otimes I_q$: $\mathrm{Tr}(P(E_0 \otimes I_q))/k \le 1/n + \frac{3}{2}\tau$. Since $w(U) = W \otimes I_q$ and $(W - I)^*(W - I) \le 4E_0$,

$$\|(w(U) - I)P\|_{\mathrm{HS}}^2/k = \mathrm{Tr}\big(P((W - I)^*(W - I) \otimes I_q)P\big)/k \le 4(1/n + \tfrac{3}{2}\tau).$$

The word w has eight letters, so $\|w(Q) - w(U)\|_{\mathrm{op}} \le 24\tau$, and in the normalized norm of the range of P, $\|w(Q)|_P - I\|_{2,k} \le 2\sqrt{1/n + 3\tau/2} + 24\tau$. On the other hand the normalized norm is unitarily invariant, so evaluating the fixed derivation of J on the restricted unitaries gives $2 = \|Q_J|_P - I\|_{2,k} \le a\|w(Q)|_P - I\|_{2,k} + bC\tau$. The two bounds give the inequality. If $n \ge 8a^2$ and $(24a + bC)\tau \le 1$, it forces $1/n + \frac{3}{2}\tau \ge 1/(4a^2)$, hence $\tau \ge 1/(12a^2)$. □

The bound holds for every number q of copies and survives perturbations of the copied generators that tend to zero in operator norm. It does not cover representation-changing amplifications, a new Baumslag–Solitar tuple, or extensions in which J is not in the normal closure of the old witness.

Example F.4 (Thin Baumslag–Solitar models)**.** Let $n \ge 3$ be odd, $q = 3^n - 2^n$, $h = (q-1)/3$ and $a_j = 3^j 2^{n-1-j}$ for $0 \le j < n$. On $\mathbb{C}^{2n}$ with basis $e_{r,j}$, $r \in \{1, 2\}$, $j \in \mathbb{Z}/n\mathbb{Z}$, put $Ue_{r,j} = \exp(2\pi i c_r a_j/q)e_{r,j}$ with $c_1 = 1$ and $(c_2 - c_1)a_0 \equiv h \pmod q$, and $V = RV_0$, where $V_0 e_{r,j} = e_{r,j-1}$ and R is the real Hadamard matrix on $E_0 = \mathrm{span}\{e_{1,0}, e_{2,0}\}$ and the identity elsewhere. Since $2a_{j+1} = 3a_j$ and $2a_0 = 2^n \equiv 3^n = 3a_{n-1} \pmod q$, $V_0 U^2 V_0^* = U^3$ exactly, so $VU^2V^* = RU^3R^*$. On E_0 the two eigenvalues of U^3 differ by the factor $\exp(2\pi i \cdot 3h/q) = \exp(-2\pi i/q)$, so $\|VU^2V^* - U^3\|_{\mathrm{op}} = \sqrt{2}\sin(\pi/q)$: the error is exponentially small in the dimension $2n$. The projections $E_j = \mathrm{span}\{e_{1,j}, e_{2,j}\}$ are permuted cyclically by V. The unitary $VUV^* = RV_0UV_0^*R^*$ preserves E_0 and is diagonal off E_0, so $W = [U, VUV^*]$ is the identity off E_0 and $(W - I)^*(W - I) \le 4E_0$; and $W \ne I$, because on E_0 the eigenvalues of U differ by $\exp(2\pi i h/q) \ne 1$ and VUV^* has the nonzero off-diagonal entry $\frac{1}{2}(\nu_1 - \nu_2)$, where $\nu_2/\nu_1 = \exp(2\pi i(c_2 - c_1)a_1/q) = -\exp(-\pi i/q)$. By Proposition F.3, no fixed extension in which the central marker lies in the normal closure of w turns copies of these models into negative models at vanishing error; and by Proposition F.1, neither does identifying the marker with W.

F.3 Locally stable groups

Let Γ be a group with a finite generating set S. It is *locally Schatten-p stable*, $1 \le p < \infty$, if for every $a > 0$ there are $\delta(a) > 0$ and a finite set R_a of words in S trivial in Γ such that every unitary assignment $U : S \to \mathrm{U}(D)$, in any dimension D, with $\max_{r \in R_a} \|r(U) - I\|_p < \delta(a)$ in the unnormalized Schatten norm, satisfies $\max_s \|U_s - \pi(s)\|_p < a$ for some homomorphism $\pi : \Gamma \to \mathrm{U}(D)$. For $p = 2$ this is Frobenius stability [47, 5]: if every asymptotic homomorphism of Γ is close to a homomorphism, the local form follows by the standard compactness argument, choosing R_a among longer and longer finite lists of relations.

Theorem F.5 (Stable sources)**.** *Let Γ be locally Schatten-p stable, with a central involution $J \in S$ that every finite-dimensional unitary representation sends to I. Let R be a finite set of words trivial in Γ containing J^2, $[J, s]$ for $s \in S$, and R_1, with largest length L. Every unitary assignment on $\mathbb{C}^D$ with $\max_{r \in R} \|r(U) - I\|_{\mathrm{op}} \le \eta$ and $\|U_J - I\|_{\mathrm{op}} > \eta$ has*

$$D \;\ge\; \Big(\frac{\delta(1)}{(1+3L)\eta}\Big)^p, \qquad \log D \ge p \log(1/\eta) - O(1).$$

Proof. Lemma 17.7 gives unitaries V_s on $\mathbb{C}^k$, $1 \le k \le D$, with $V_J = -I_k$ and $\|r(V) - I\|_{\mathrm{op}} \le (1+3L)\eta$ for $r \in R$, hence $\|r(V) - I\|_p \le k^{1/p}(1+3L)\eta$. If this were below $\delta(1)$ for all $r \in R_1$, there would be a representation π of Γ on $\mathbb{C}^k$ with $\|V_J - \pi(J)\|_p < 1$; but $\pi(J) = I$ and $\| - I_k - I_k\|_p = 2k^{1/p} \ge 2$. So $k^{1/p}(1+3L)\eta \ge \delta(1)$. □

So no locally Schatten-stable group satisfies the hypothesis $\log D_n = o(\log(1/\eta_n))$ of Proposition 17.8. The theorem does not say that all approximate models are close to exact ones; it prices separated models polynomially.

Example F.6 (The higher-rank Deligne extension)**.** Let E be the preimage of $\mathrm{Sp}(8, \mathbb{Z})$ in the universal cover of $\mathrm{Sp}(8, \mathbb{R})$, let z generate the infinite cyclic kernel of $E \to \mathrm{Sp}(8, \mathbb{Z})$, and put $Q = E/\langle z^4 \rangle$ and $J = z^2\langle z^4 \rangle$, a central involution with $J \ne 1$. Deligne's theorem [50], stated there for $\mathrm{Sp}(2n, \mathbb{Z})$ with $n \ge 2$, says that every finite-index subgroup of E contains z^2; see also [15]. A finite-dimensional unitary representation of the finitely generated group Q has a finitely generated linear image, which is residually finite by Mal'cev's theorem [93]; if it did not kill J, some finite quotient of Q, hence of E, would not kill z^2. So every finite-dimensional unitary representation sends J to I. The group E has property $[T_2]$, since $\mathrm{Sp}(8, \mathbb{R})$ is simple of real rank four [5, Theorem 2.10]; groups with vanishing second cohomology in every unitary representation are Frobenius stable, by the theorem of De Chiffre, Glebsky, Lubotzky and Thom [47] in the form of [5, Theorem 1.1(F)]; and $\langle z^4 \rangle$ lies in the intersection of the finite-index subgroups of E, so the quotient Q is Frobenius stable [5, Lemma 2.12]. By Theorem F.5 with $p = 2$, every operator-norm model of Q with J separated has $D \ge c_Q \eta^{-2}$. These are imported results; their statements have been checked against the primary sources, not their proofs.

F.4 The dyadic pattern and Abels' group

Corollary F.7. *Let $\langle S \mid R \rangle$ contain generators w, x, y, J and the relators J^2, $[J, s]$ for $s \in S$, $wyw^{-1}y^{-1}J^{-1}$ and $xyx^{-1}y^{-2}$, with largest relator length L. Every unitary assignment on $\mathbb{C}^D$ with operator-norm relator errors at most η and $\|U_J - I\|_{\mathrm{op}} > \eta$ has $D \ge 1/(2(1+3L)\eta)$.*

Proof. Lemma 17.7 gives V_s on $\mathbb{C}^k$, $k \le D$, with $V_J = -I$ and relator errors at most $C\eta$, $C = 1 + 3L$. The first relator gives $\|V_w V_y V_w^* V_y^* + I\|_{\mathrm{op}} \le C\eta$, hence $\|V_w V_y V_w^* + V_y\|_{\mathrm{op}} \le C\eta$; the second gives $\|V_x V_y V_x^* - V_y^2\|_{\mathrm{op}} \le C\eta$. Proposition 17.5 gives $k \ge 1/(2C\eta)$. □

The Abels quotient B of Section 17.2 has these words, so its separated operator-norm models need $\log D \geq \log(1/\eta) - O(1)$ and it cannot meet the criterion of Proposition 17.8; this is consistent with Corollary 17.6, which proves the floor for its game. The same holds for every source in which the marker is in the normal closure of a dyadic sign-and-dilation pattern by a fixed derivation, since a fixed derivation changes the errors only by a constant factor.

F.5 Sign and dilation in the limits of Theorem A

The class theorem needs the sign and the dilation on the original operators, with controlled cost. In the limits of the proof of Theorem A for a game without a finite-dimensional perfect strategy they are always present, the sign exactly and the dilation to every accuracy.

Proposition F.8. *Let $\mathcal{M}$ be a finite von Neumann algebra with a faithful normal tracial state and no type I central summand, for instance a limit $\mathcal{M}$ of the proof of Theorem 7.2 for a game without a finite-dimensional perfect strategy. For every $n \geq 1$, $\mathcal{M}$ contains unitaries X, Y and a reflection T with $TYT = -Y$ and $\|XYX^* - Y^2\|_{\mathrm{op}} = 2\sin(\pi/2^n)$.*

Proof. $\mathcal{M}$ contains a unital copy of $M_d(\mathbb{C})$ for every d [27, Lemma 1.4]. In M_d, $d = 2^n$, put $Y = \mathrm{diag}(z_j)$ with $z_j = e^{2\pi i j/d}$, let T exchange e_j and $e_{j+d/2}$, and let X be the permutation unitary with $Xe_{\varpi(j)} = e_j$, $\varpi(j) = 2(j \bmod d/2) + \lfloor 2j/d \rfloor$. Then $TYT = -Y$ since $z_{j+d/2} = -z_j$, and XYX^*Y^{-2} is I on the first half of the coordinates and $e^{2\pi i/d}I$ on the second, as in [101, Example 11.11]. □

So the limit algebra cannot distinguish the class of Theorem 17.3. Bounded polynomial approximation recovers these operators from the tested generators to any fixed accuracy, for each perfect trace, but gives no bound on the degree and coefficients of the polynomials as the accuracy decreases, and no uniform neighbourhood of the trace. The class theorem has these data because its operators are actual local operators on the original state, available before any limit is taken. The next proposition shows how large the missing bound can be.

F.6 Slow recovery from four projections

For a word polynomial $F = \sum_w c_w w$ in finitely many letters put $B(F) = \sum_w |c_w||w|$. This is the quantity in the actual-mirror transport estimate (6.1): if every tested generator has mirror error δ, then F has mirror error at most $B(F)\delta$. Let $\mathcal{A} = \bigotimes_{j\geq 1} M_2$ be the dyadic UHF algebra with its product trace, Z_j, X_j the Pauli reflections at site j, $\alpha_j = 2^{-2^j}$, $a = \sum_j \alpha_j Z_j$, $b = \sum_j \alpha_j X_j$, $U = e^{ia}$ and $V = e^{ib}$. In $\widetilde{\mathcal{A}} = M_2(\mathcal{A})$ let

$$S_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad S_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & U \\ U^* & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 0 & V \\ V^* & 0 \end{pmatrix}, \quad P_\ell = \tfrac{1}{2}(1 + S_\ell).$$

Proposition F.9 (Slow recovery)**.** *The four projections P_ℓ generate $\widetilde{\mathcal{A}}$, which has no nonzero finite-dimensional representation and whose closure in its faithful trace is the hyperfinite II_1 factor. For $r \geq 3$ let $F_1, \ldots, F_{2r}$ be word polynomials in four letters that are formally self-adjoint, $F_j = F_j^*$, with $\|F_j(P)\|_{\mathrm{op}} \leq 2$ and, in the normalized trace norm $\|\cdot\|_2$,*

$$\|F_j(P)^2 - 1\|_2 \leq \kappa_r, \qquad \|F_j(P)F_k(P) - s_{jk}F_k(P)F_j(P)\|_2 \leq \kappa_r, \qquad \kappa_r = 10^{-6} r^{-5/2},$$

where s_{jk} are the signs of the standard r-qubit Pauli array of Theorem 6.4. Then $\max_j B(F_j) > 5\cdot 10^{-7}\, r^{-5/2}\, 2^{2^{r-1}}$. A polynomial whose value at P is self-adjoint can be replaced by $\frac{1}{2}(F + F^*)$, which has the same value and no larger B.

Proof. Generation. Since $\|a\|_{\mathrm{op}}, \|b\|_{\mathrm{op}} < 1$, the continuous logarithm recovers a, b from U, V. The Z_j commute and $\alpha_j > \sum_{k>j} \alpha_k$, so the map $\sigma \mapsto \sum_j \alpha_j \sigma_j$ is a homeomorphism of $\{\pm 1\}^{\mathbb{N}}$ onto the spectrum of a, and every Z_j is a continuous function of a. Next $[Z_j, [Z_j, b]] = 4\alpha_j X_j$, so $X_j \in C^*(a, b)$, and the Z_j, X_j generate $\mathcal{A}$. In $\widetilde{\mathcal{A}}$, S_0 and S_1 give the matrix units of M_2, and the corners of S_2, S_3 give U, V. A nonzero finite-dimensional representation would restrict to unital representations of the unital subalgebras M_{2^k} for all k in one fixed dimension, which is impossible. The trace is the product trace, whose closure is the hyperfinite II_1 factor.

Truncation. Let a_n, b_n, U_n, V_n and $P_\ell^{(n)}$ be obtained by stopping the sums at site n; they lie in $M_2 \otimes M_{2^n} = M_{D_n}$, $D_n = 2^{n+1}$, a unital subalgebra on which the trace is the normalized matrix trace. With $\tau_n = \sum_{j>n} \alpha_j \le 2 \cdot 2^{-2^{n+1}}$, $\|e^{ia} - e^{ia_n}\|_{\mathrm{op}} \le \|a - a_n\|_{\mathrm{op}} \le \tau_n$, so $\|P_\ell - P_\ell^{(n)}\|_{\mathrm{op}} \le \tau_n$; telescoping each word, whose letters are contractions, gives $\|F(P) - F(P^{(n)})\|_{\mathrm{op}} \le B(F)\tau_n$.

Contradiction. Take $n = r - 2$, so $D_n = 2^{r-1}$, and suppose $t = \max_j B(F_j)\tau_n \le \kappa_r$. The matrices $H_j = F_j(P^{(n)})$ are self-adjoint, because F_j is formally self-adjoint and the $P_\ell^{(n)}$ are self-adjoint, and $\|H_j\|_{\mathrm{op}} \le 3$. Replacing $F_j(P)$ by H_j changes the square by at most $5t$ and each ordered product by at most $5t$, so $\|H_j^2 - 1\|_2 \le 6\kappa_r$ and $\|H_j H_k - s_{jk} H_k H_j\|_2 \le 11\kappa_r$. Let $Q_j = \operatorname{sgn}(H_j)$, with $\operatorname{sgn} 0 = 1$; the scalar inequality $|\operatorname{sgn} x - x| \le |x^2 - 1|$ gives $\|Q_j - H_j\|_2 \le 6\kappa_r$, and $\|Q_j Q_k - H_j H_k\|_2 \le \|(Q_j - H_j)Q_k\|_2 + \|H_j(Q_k - H_k)\|_2 \le 24\kappa_r$, so the reflections Q_j have signed pair defects at most $59\kappa_r$. Let Bob use the Q_j and Alice their transposes on Φ_{D_n}: the mirrors are exact and both parties have these defects. Lemma 6.5, with $m = 2r$, $C_r = 2r(1 + 4r) \le 10r^2$ and $p = (2C_r + 1)^2 (59\kappa_r)^2/4$, gives an output Φ_{2^r} with singlet infidelity at most $2rp \le \frac{441}{2} \cdot 59^2 \cdot 10^{-12} < 10^{-6}$. But product local isometries applied to Φ_{D_n} give singlet fraction at most $D_n/2^r = \frac{1}{2}$ with Φ_{2^r}, by the Schmidt-rank bound of Proposition F.11. So $\max_j B(F_j) > \kappa_r/\tau_n \ge 5 \cdot 10^{-7} r^{-5/2} 2^{2^{r-1}}$. □

The accuracy κ_r is inverse-polynomial in r, of the order of the fixed-fidelity tolerance in the proof of Theorem 7.2. So no proof that uses only finitely many tested projections, absence of finite-dimensional representations and a type II tracial closure can bound the transport budget of its Pauli polynomials by e^{Cr}: here the budget is doubly exponential, at one faithful trace, in a nuclear and hyperfinite example. This is not a counterexample to the floor: no finite game presenting this tuple, and no retained-state upper strategy, is constructed, and other decoders might avoid these polynomials.

F.7 Single-copy thresholds

For a finite-dimensional state ρ_{AB} let $f_r(\rho_{AB})$ be the largest overlap $\langle \Phi_{2^r} | \sigma | \Phi_{2^r} \rangle$ of a state σ obtained from ρ_{AB} by product local isometries followed by discarding junk. For a game G in class (S) or (C), with quantum value one and no finite-dimensional perfect strategy, and $0 < \gamma < \frac{1}{2}$, let $\Theta_G(r; \gamma)$ be the supremum of the losses $\varepsilon \le 1$ such that every finite-dimensional strategy of loss at most ε has $f_r \ge 1 - \gamma$; Theorem A says that this threshold is positive.

Proposition F.10 (Exponential thresholds suffice)**.** *If $\Theta_G(r; \gamma) \ge e^{-C_G r}$ for all $r \ge 1$, then $d_G^{\rightarrow}(\varepsilon) \ge \frac{1-2\gamma}{C_G} \ln(1/\varepsilon) - O_{G,\gamma}(1)$.*

Proof. Given ε, let $r = \lfloor \ln(1/\varepsilon)/C_G \rfloor - 1$. If $r \ge 1$, then $\varepsilon < e^{-C_G r} \le \Theta_G(r; \gamma)$, so every strategy of loss at most ε has $f_r \ge 1 - \gamma$, and (2.4) with $D = 2^r$ gives $E_{\mathrm{D}}^{A \to B} \ge (1 - 2\gamma) r - h(\gamma)$; local isometries and discarding do not increase the directional rates. □

The converse cannot be derived from the definitions: a large distillable entanglement does not give a good single-copy extraction at a target logarithmic in it.

Proposition F.11 (A separation)**.** *(a) If a unit vector has Schmidt rank at most s, then every product local isometry followed by discarding junk gives overlap at most s/K with Φ_K.*

(b) For $n = 2^t$ let $\psi_n = \frac{1}{2}|00\rangle \otimes \Phi_n + \frac{\sqrt{3}}{2}|11\rangle \otimes \Phi_{2^n}$, with orthogonal local flags. Then $E_{\mathrm{D}}^{A\to B}(\psi_n) = h(\frac{1}{4}) + \frac{t}{4} + \frac{3n}{4}$; for every fixed r and $t \geq r$, $f_r(\psi_n) = 1$; but at $r_n = 2t$,

$$f_{r_n}(\psi_n) \;\leq\; \Big(\frac{1}{2\sqrt{n}} + \frac{\sqrt{3}}{2}\Big)^2 \;\longrightarrow\; \frac{3}{4}.$$

So for every fixed $\gamma < \frac{1}{4}$ and large n, single-copy extraction of $r_n = 2\log n$ pairs at fidelity $1-\gamma$ fails, although the distillable entanglement is of order $n = 2^{r_n/2}$.

Proof. (a) Product local isometries do not increase the Schmidt rank. Discarding junk systems A_j, B_j from a vector $\sum_{k\leq s}\sqrt{\lambda_k}\, a_k \otimes b_k$ leaves the mixture of the vectors $\sum_k \sqrt{\lambda_k}\langle\alpha|a_k\rangle \otimes \langle\beta|b_k\rangle$ over basis vectors α, β of the junk, each of Schmidt rank at most s. For a vector $v = \sum_{k\leq s} c_k x_k \otimes y_k$ with orthonormal $(x_k), (y_k)$, $|\langle \Phi_K, v\rangle| \leq K^{-1/2}\sum_k |c_k| \leq \sqrt{s/K}\,\|v\|$ by Cauchy–Schwarz; summing over the mixture gives s/K.

(b) ψ_n is pure with Schmidt coefficients $\frac{1}{4n}$ (n times) and $\frac{3}{4\cdot 2^n}$ (2^n times), so its directional rates are its entanglement entropy [12], as displayed. For $t \geq r$ both n and 2^n are multiples of 2^r, and controlled local isometries, reading the flags, factor Φ_n and Φ_{2^n} as Φ_{2^r} times a maximally entangled junk exactly. At $r_n = 2t$, $K = n^2$; for product local isometries V and the projection Π onto Φ_K times the junk, $\|\Pi V\psi_n\| \leq \frac{1}{2}\|\Pi V(|00\rangle\Phi_n)\| + \frac{\sqrt{3}}{2} \leq \frac{1}{2}\sqrt{n/K} + \frac{\sqrt{3}}{2}$ by (a), and f_{r_n} is the supremum of $\|\Pi V\psi_n\|^2$. □

No game's loss is identified with these states, so the proposition is not a counterexample to any game-specific converse: it shows that a converse needs information about the game. Extraction fails here only for target infidelities below about $\frac{1}{4}$.

F.8 Continuous amplification, algebraic augmentation and kernel leakage

Proposition F.12 (Continuous amplification after marker compression)**.** *Keep exactly the thin tuple U_n, V_n of Example F.4 in $K_n = U(2)^n \rtimes C_n$, for odd $n \geq 3$. Fix a finite presentation containing u, v, J, the BS relator $vu^2v^{-1}u^{-3}$, J^2 and all centrality relators $[J, s]$. Fix a free derivation of J as a product of $a \geq 1$ conjugates of $w^{\pm 1}$, $w = [u, vuv^{-1}]$, and b relator conjugates; a, b and the derivation do not vary with n.*

For any continuous finite-dimensional unitary representation π of K_n, assign $u = \pi(U_n)$, $v = \pi(V_n)$ and arbitrary other unitaries. Suppose the maximal operator-norm defining defect η tends to zero and $\|T_J - I\|_{\mathrm{op}} > \eta$. Let k be the negative rank of the nearest-sign rounding of T_J. Then

$$\log_2 k \geq \frac{\log_2(1/\eta)}{3a^2\log_2 3} - O_{\mathcal{P}}(1).$$

The bounded term may depend on the fixed derivation. Multiplicities, highest weights and ambient dimensions may vary arbitrarily. Continuity and exactness of the group representation are hypotheses; this is not a statement about arbitrary nearly represented tuples.

Proof. Round T_J to H and put $P = (I - H)/2$. For small η, $\|H - T_J\|_{\mathrm{op}} \leq \eta$, $\|[P, T_s]\|_{\mathrm{op}} \leq 3\eta/2$, and $P \neq 0$. Block-polar repair as in Lemma 17.7 gives Q_s commuting with P, $\|Q_s - T_s\|_{\mathrm{op}} \leq 3\eta$ and defining defects at most $C\eta$, $C = 1 + 3L_0$, where L_0 is the maximal defining-relator length. Telescope the eight-letter w and evaluate the fixed derivation in P's normalized Hilbert–Schmidt norm:

$$2 \leq a\|(w(T) - I)P\|_{HS}/\sqrt{k} + (24a + bC)\eta.$$

Put $q = 3^n - 2^n$. The local BS defect is in one $SU(2)$ block with eigenangles $\pm\theta$, where $2\sin(\theta/2) = \sqrt{2}\sin(\pi/q)$ and $\theta \geq 2\sqrt{2}/q$. A local irreducible of highest weight ℓ has weights $-\ell, -\ell+2, \ldots, \ell$. If $\ell\theta \geq \pi/2$, one weight angle lies in $[\pi/2, \pi/2 + 2\theta]$, forcing defect at least $\sqrt{2}$, impossible for $\eta < 1/2$. Otherwise the sine bound gives $\ell \leq \pi\eta q/(4\sqrt{2}) < \eta q$. Cyclic conjugation gives this at every block. Thus a local $g \in SU(2)$ satisfies $\|\pi(g) - I\|_{\mathrm{op}} \leq (\pi/2)\eta q\|g - I\|_{\mathrm{op}}$.

The exact local calculation in the thin tuple gives an order-three diagonal D_0 with $\|W - D_0\|_{\mathrm{op}} \leq 10\pi/(3q)$, $W = [U_n, V_n U_n V_n^*]$. Hence $\|\pi(W) - \pi(D_0)\|_{\mathrm{op}} < 17\eta$. Set $G_j = \pi(V_n^j D_0 V_n^{-j})$ and $B_j = \pi(V_n^{j+1} U_n V_n^{-j-1})$, $0 \leq j < n$. The G_j commute, $G_j^3 = I$, $[B_j, G_l] = 0$ for $l > j$, and $\|B_j^* G_j B_j - G_j^*\|_{\mathrm{op}} < 10\eta$. Indeed the local near-Pauli-X conjugator has $SU(2)$ error at most $2\pi/q$, giving $\pi^2\eta < 10\eta$. Word transport gives

$$\|[P, G_j]\|_{\mathrm{op}} \leq (46 + 3j)\eta, \qquad \|[P, B_j]\|_{\mathrm{op}} \leq (4.5 + 3j)\eta.$$

In P/k, put $p_j = \Pr(G_j \neq 1)$. The marker inequality gives $p_0 \geq 1/(3a^2)$ eventually; approximate invariance under $\pi(V_n)$ gives $p_j \geq 1/(3a^2) - 3j\eta/2$. The marker also forces a nontrivial local highest weight, so $q > 1/\eta$ and $m = \lfloor \log_3(1/\eta) \rfloor < n$ is available.

Jointly measure $G_0, \ldots, G_{m-1}$, obtaining ternary X_j. Conjugation by B_j approximately flips the $\pm$ labels and preserves all later labels; its total-variation error is $t_j \leq 25m\eta$. Averaging with that law preserves the tail marginal and activity and makes active signs equiprobable conditional on the tail. The three-valued conditional-entropy estimate of Alhejji and Smith [2] gives

$$H(X_j \mid X_{j+1}, \ldots) \geq p_j - h(t_j/2) - t_j/2.$$

There is no exponential dependence on the number of tail labels. Summing gives $H(X) \geq m/(3a^2) - (3/4)m^2\eta - m[h(12.5m\eta) + 12.5m\eta]$.

Joint pinching of P/k has outside-P mass $\alpha \leq (46m^2\eta)^2 = \alpha_*$ and rank at most $3^m k$. Pinch again across P and its complement: its entropy is at most $\log k + h(\alpha) + \alpha m \log_2 3$, and is at least $H(X)$. In the small-parameter monotone range,

$$\log k \geq \frac{m}{3a^2} - \frac{3}{4}m^2\eta - m[h(12.5m\eta) + 12.5m\eta] - h(\alpha_*) - \alpha_* m \log_2 3.$$

All error terms tend to zero for the chosen m, proving the negative-sector rank bound. At the certified loss $\varepsilon = C\eta^2$ of Proposition 17.8, the particular upper state Φ_{12k} therefore costs a positive multiple of $\log(1/\varepsilon)$ up to constants. This does not lower-bound every compiled-game strategy, nor compare with a smaller actual loss or independently improved precision. □

Proposition F.13 (Spectral multiplicity obstruction)**.** *Let a finitely presented group have an explicitly central involution J and no nonzero finite-dimensional representation with $J = -I$. Suppose J lies in the normal closure of w, with a fixed derivation using a conjugates of $w^{\pm 1}$ and b relator conjugates. Any separated operator-norm models with $\eta_n \to 0$ and $\|T_J - I\|_{\mathrm{op}} > \eta_n$ satisfy*

$$\operatorname{rank} 1_{[t,\infty)}(|w(T_n) - I|) \longrightarrow \infty \quad (0 < t < 2/a).$$

In particular, a uniformly bounded unnormalized Schatten-p norm of $w(T_n) - I$, for any finite p, prevents this route.

Proof. Round J to H with negative rank k. Replace the eigenvalues of $W = w(T)$ within distance t of 1 by 1, obtaining W_0 with $\|W - W_0\|_{\mathrm{op}} \leq t$ and $\operatorname{rank}(W_0 - I) = N_t(W)$. Substitute W_0 in the derivation and omit relator factors. The product A_0 has $\operatorname{rank}(A_0 - I) \leq aN_t(W)$ and $\|H - A_0\|_{\mathrm{op}} \leq (b+1)\eta + at < 2$ eventually. If $k > aN_t(W)$, a vector in the negative-H space and $\ker(A_0 - I)$ contradicts this. Thus $1 \leq k \leq aN_t(W)$. Bounded multiplicity along a

subsequence would give bounded nonzero compressed dimensions. Block-polar repair makes their defects tend to zero; a fixed-dimensional compact subsequence gives an exact forbidden representation. Finally $t^p N_t(W) \le \|W - I\|_p^p$. □

Proposition F.14 (Polynomial identities of the slow-recovery tuple)**.** *For positive summable α_j with infinitely many distinct values, form $U = \exp(i\sum \alpha_j Z_j)$, $V = \exp(i\sum \alpha_j X_j)$, and the four P_ℓ as in Proposition F.9. Every noncommutative $*$-polynomial equation of this tuple, with arbitrary complex coefficients, holds simultaneously at*

$$P_0^{(0)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad P_1^{(0)} = P_2^{(0)} = P_3^{(0)} = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

This includes $\alpha_j = 2^{-2^j}$. Finitely many explicit polynomial auxiliaries also transfer, including their projection, reflection or unitary domain equations. Independently introduced implicit variables, nonpolynomial spectral functions and arbitrary norm inequalities are not covered. A perfect polynomial realization in the faithful product trace therefore cannot define a game without a finite perfect strategy.

Proof. Put $A(z) = zE_+ + z^{-1}E_-$ and $B(z) = zF_+ + z^{-1}F_-$. Every nontrivial cyclically reduced free word has nonconstant Laurent-polynomial trace. For a mixed alternating word, its coefficient at exponent $\sum |n_i| + |m_i|$ is the cyclic product of nonzero overlaps between Z and X eigenvectors, hence is nonzero; a pure power is immediate. Cyclic reduction preserves trace. Determinant one makes a scalar word value $\pm I$, so it occurs at only finitely many nonzero z. Distinct free words are therefore noncollinear at infinitely many distinct sites $z_j = e^{i\alpha_j}$.

Summability gives norm-convergent infinite tensor products. In a finite linear dependence of word values, fix one word and choose for each competitor a distinct site where their local values are noncollinear. Slice there by a functional killing the competitor but not the chosen word. Their finite tensor product kills every other word and leaves a nonzero scalar times a unitary, so the chosen coefficient is zero. Thus $\mathbb{C}[F_2] \to UHF$ is algebraically injective.

The four projections are formal elements of $M_2(\mathbb{C}[F_2])$. Injectivity makes every polynomial relation of the concrete tuple formal. Apply augmentation $u, v \mapsto 1$ to obtain the displayed tuple, substituting any explicit polynomial auxiliaries first. For a perfect game in the faithful trace, a supported losing product has $\tau(QR) = \tau((QR)^*(QR)) = 0$, hence $QR = 0$. Partitions, context commutations and forbidden products transfer. The augmented projections and transposes on Φ_2 give a perfect synchronous strategy; for a constraint-variable game use the joint PVM in each commuting augmented context. Discontinuity of the algebraic augmentation in the C^* norm is allowed and does not contradict the completed UHF algebra's absence of finite-dimensional representations. □

Proposition F.15 (An original-state bound conditional on leakage)**.** *Let U_s be actual Bob source unitaries and W_s actual Alice unitary mirrors on a purification of the original ρ, with summed squared mirror errors at most $C_0\varepsilon$. Let fixed bounded residuals R_i, including $U_J + I$, satisfy $Q = \sum_i \|R_i(U)\psi\|^2 \le C_1\varepsilon$ and $\sum_i \|R_i(U)\|_{\mathrm{op}}^2 \le K^2$. Assume every irreducible d-dimensional unitary tuple V satisfies*

$$d^{-1}\sum_i \|R_i(V)\|_{HS}^2 \ge A d^{-\varkappa}, \qquad A, \varkappa > 0.$$

Apply the Alice-local recovery of Theorem 6.10 to obtain σ on $\overline{B} : B$. Put

$$H_U = \sum_s (U_s^T \otimes I - I \otimes U_s)^*(U_s^T \otimes I - I \otimes U_s), \quad P_0 = 1_{\{0\}}(H_U), \quad q = \operatorname{Tr}\sigma(I - P_0).$$

Then $\operatorname{Tr}\sigma H_U \le 2C_0\varepsilon$. *For* $q<1$ *set* $Q_* = (\sqrt{Q}+K\sqrt{q})^2/(1-q)$ *and* $u = Q_* M^{-\varkappa}/A + \sqrt{q}$. *Every integer* $M \ge 2$ *with* $u \le 1/2$ *satisfies*

$$E_{\mathrm{D}}^{A\to B}(\rho) \ge (1-2u)\log M - h(u).$$

Proof. Recovery is an existing Alice-local channel, so cannot increase the rate; Bob's marginal and residual expectations are unchanged. Compare σ with $\tau = P_0\sigma P_0/(1-q)$ at half trace distance at most $\sqrt{q}$. Applying the residual row operator to the projected purification bounds total defect by Q_*, without commuting with P_0.

Decompose Bob's tuple algebra into M_{d_λ} tensor identity on multiplicities. Vectorizing the kernel identifies the commutant, giving Φ_{d_λ} tensor multiplicity spaces with common local labels. Residuals act only on irreducible factors, so $Q_* \ge A\sum p_\lambda d_\lambda^{-\varkappa}$. Sectors with $d_\lambda < M$ have mass at most $Q_* M^{-\varkappa}/A$. Discard multiplicities and convert every large Φ_d to Φ_M by the subset instrument in Proposition 11.24, using default product outputs elsewhere [109]. The ideal singlet fraction is at least $1 - Q_* M^{-\varkappa}/A$; on σ it is at least $1-u$ by contraction. Isotropic twirling and hashing give $\log M - h(u) - u\log(M^2-1)$, at least the claimed bound. The subset instrument, local discards, twirling and hashing form the stated one-way protocol. □

For $z = \varepsilon + q \to 0$, $Q_* \le Cz$. Choose $M = \lfloor [A/(Cz\log(1/z))]^{1/\varkappa} \rfloor$, eventually at least two. Then $u \le 1/\log(1/z) + \sqrt{z}$ and

$$E_{\mathrm{D}}^{A\to B}(\rho) \ge \varkappa^{-1}\log\frac{1}{\varepsilon+q} - \varkappa^{-1}\log\log\frac{1}{\varepsilon+q} - O(1).$$

In particular $q \le C'\varepsilon^a$, $a>0$ fixed, gives coefficient $\min(1,a)/\varkappa$ with $O(\log\log(1/\varepsilon))$ loss. No such leakage estimate is asserted for an arbitrary game.

Proposition F.16 (Locally Schatten-stable sources)**.** *Let a source with central involution* J *be locally stable in the unnormalized Schatten-*p *norm,* $1 \le p < \infty$, *in the sense of Theorem F.5, and have no nonzero finite-dimensional representation with* $J = -I$. *Include a finite stability relation list for tolerance one and marker residual* $V_J + I$. *Every* d*-dimensional unitary tuple satisfies*

$$d^{-1}\sum_i \|R_i(V)\|_{HS}^2 \ge b^2 d^{-\varkappa}, \qquad b = \min\{\delta(1),1\}, \quad \varkappa = \max\{1, 2/p\}.$$

Proof. Every finite-dimensional representation kills J, since otherwise its negative eigenspace supplies a forbidden representation. If all listed residuals had Schatten-p norm below b, stability gives $\|V_J - I\|_p < 1$, while $\|V_J + I\|_p < 1$. Their sum is below two, contradicting $\|2I_d\|_p = 2d^{1/p} \ge 2$. Thus one residual has norm at least b. For $p \le 2$ use $\|X\|_p \le d^{1/p-1/2}\|X\|_{HS}$; for $p \ge 2$, use $\|X\|_p \le \|X\|_{HS}$. Square and divide by d. □

The exactly mirrored coefficient in Proposition F.15 is therefore $\min(1, p/2)$. Exact mixing of mirrored sectors, with arbitrary multiplicities and sector coherence, cannot create a sublogarithmic family under these source assumptions. Fixed relator derivations and actual mirrors transfer added stability words with $O(\sqrt{\varepsilon})$ constants if the tested presentation differs. Neither normalized Hilbert–Schmidt stability nor property (T) alone substitutes for this hypothesis; see [47, 91] for related stability and nonapproximability results.

Proposition F.17 (Necessary populated slow modes)**.** *Under Proposition F.15's fixed-source tracial-defect hypothesis, if* $\varepsilon_n \to 0$ *and* $E_{\mathrm{D}}^{A\to B}(\rho_n)/L_n \to 0$, *then, for every fixed* $a>0$, $q_n > \varepsilon_n^a$ *eventually. Equivalently* $\log(1/q_n)/L_n \to 0$. *If* $\beta_0 = 2C_0 > 0$, *then*

$$\operatorname{Tr}[\sigma_n 1_{(0,\,2\beta_0\varepsilon_n/q_n]}(H_n)] \ge q_n/2.$$

Thus the state must put $\varepsilon_n^{o(1)}$ *weight on nonzero eigenmodes at most* $\varepsilon_n^{1-o(1)}$. *An unused block with a tiny gap is insufficient.*

Proof. An infinite subsequence with $q_n \le \varepsilon_n^a$ contradicts the logarithmic bound on that subsequence. Taking logarithms gives equivalence. Since $\operatorname{Tr}\sigma_n H_n \le \beta_0\varepsilon_n$, spectral Markov bounds mass above $2\beta_0\varepsilon_n/q_n$ by $q_n/2$, leaving the claimed mass. If $C_0 = 0$, then $q = 0$ and the preceding bound already excludes a sublogarithmic sequence; no positive-threshold expression is needed.

If the smallest positive eigenvalue is at least $c\varepsilon_n^\beta$, $0 \le \beta < 1$, then $q_n \le C\varepsilon_n^{1-\beta}$ and the rate is at least $(1-\beta)L_n/\varkappa - (\log L_n)/\varkappa - O(1)$. Replace U_s by V_s with $\max_s\|V_s - U_s\|_{\mathrm{op}} \le L_0\sqrt{\varepsilon}$ for fixed L_0. Their transpose-mirror errors on the same recovered σ are still $O_{L_0}(\sqrt{\varepsilon})$; fixed polynomial residuals change by that order by telescoping. The whole proof applies to H_V. A sublogarithmic sequence therefore cannot admit such perturbations with polynomially small kernel leakage. A small gap alone is not asserted sufficient for a counterexample. □

Remark F.18 (Qualitative thresholds and selectable depth). Positive fixed-target thresholds alone cannot imply an exponential-in-target lower bound: 2^{-2^r} is an explicit counterexample to that numerical implication, even if the proposed exponential constant is non-effective. A sublogarithmic family can still have diverging entanglement, so fixed-target divergence does not exclude it.

For the depth bookkeeping write $\ell(e) = \ln(1/e)$. An estimate $m \ge c\ell(e)$, $\eta \le e\exp(Cm)$, with $cC < 1$, does not ensure small η without an upper control on the chosen m. A selectable $m = \lfloor a\ell(e)\rfloor$, $aC < 1$, or a comparable upper bound, is sufficient: η is a positive power of e and $m^4\eta \to 0$.

Proof. Every 2^{-2^r} is positive but it is smaller than e^{-Cr} for all large r, for any fixed C. A lower bound on m permits arbitrarily large $e\exp(Cm)$. For the selectable depth, $\eta \le e^{1-aC}$ and the logarithmic factor m^4 cannot offset its positive exponent. If the depth is instead measured with $\log_2(1/e)$, the equivalent condition is $aC < \ln 2$; the explicit natural logarithm here fixes that convention. □

G The point window: fillings and subdivision obstructions

Appendix G.1 proves Proposition 13.6a, and Appendix G.2 proves Proposition 13.9a. The remaining subsection gives the raw cross-copy obstruction.

The lower bound of Theorem 13.7 uses copies $S_i = a^{3i}sa^{-3i}$, $T_i = a^{3i+1}sa^{-3i-1}$, $0 \le i < N$, of the three-cycle pair, and bounds every raw cross-copy commutator, a far commutator of length $O(N)$, by $\eta_0 \le CN^u e$ when the far-commutator areas grow like R^u; balancing $N\kappa^2$ against $\kappa \le CN^u e$ gives the exponent $1/(2u+1)$. This appendix proves that no such worst-pair estimate holds with $u < 2$ on the original source models, even at the scale $N \asymp e^{-2/(2u+1)}$ at which the proof uses it. So this route stops at $u = 2$, that is at $\varepsilon^{-1/5}$, whatever the true growth of the areas; reaching the quarter needs a different estimate or a different extraction.

The models are the collective models of Proposition 13.9 with one particular legal choice of the spanning trees. The curl equations of that proof may be solved along any spanning tree of interior edges of each region in which the source is a leaf attached by its bottom edge: the holonomy at every plaquette depends only on the fields, and the protected supports only on the labels. We use the row-spine trees of the proof of Proposition 15.4, for the region $B = R_5$ and for T.

Lemma G.1. *In the collective model with row-spine trees, on every row $2 \le j \le M-2$ the only a-links carrying a nonzero rotation are the one at column $M-1$, with the factor $\exp(i\varphi_B Q_B)$, $|\varphi_B| \le 2\pi j/M$, and the one at column 0, with the factor $\exp(i\varphi_T Q_T)$, $|\varphi_T| \le 2\pi j/(M-1)$. For $3 \le j \le M-2$ and every clock c of row j, the block W_c of a^{4M} at c, a unitary on $(\mathbb{C}^2)^{\otimes L}$, satisfies $\|W_c - I\|_{\mathrm{op}} \le 24\pi j/M$ on the whole qubit space.*

Proof. The angles are those computed in the proof of Proposition 15.4 for the same trees: a tree edge that cuts off a subtree S not containing the source carries $|\varphi_R| = \pi|S|/|R|$, the subtree below the spine edge of row j has at most $j(M-1)$ plaquettes of B and at most $j(M-2)$ of T, and $|B| = M(M-1)/2$, $|T| \geq (M-1)(M-2)/2$; the other a-links of these rows are not tree edges, and neither exact move lies on them. Here each region contributes one rotation $\exp(i\varphi Q_R)$ with a projection Q_R, so $\|\exp(i\varphi Q_R) - I\|_{\mathrm{op}} \leq |\varphi|$ on the whole space. The generator a uses the base links of [55, §6] in the convention of the proof of Proposition 15.4, through their qubit-permutation representation, so without the rotations the M-step circuit of row j from any column is the qubit permutation of a colour permutation conjugate to the P of that proof, with $P^4 = I$ for $j \geq 3$. The product W_c of the $4M$ link factors traverses the row four times and differs from the unrotated product, which is I, in eight factors; telescoping gives $\|W_c - I\|_{\mathrm{op}} \leq 4[2\pi j/M + 2\pi j/(M-1)] \leq 24\pi j/M$. □

Theorem G.2 (The raw cross-copy obstruction)**.** *For every integer $M \geq 1152\pi$ the collective model of Proposition 13.9 with row-spine trees, on Φ_{D_M} with Alice's transposes, has all 22 relators exact except H4, $J = -I$, every signed mirror exact, precision $e_M \leq 8\pi/M^2$, directional rates $4M + 4 + 2\log M$, and*

$$\|([s, a^{4M+1}sa^{-4M-1}] - I)\Phi_{D_M}\|^2 \;\geq\; \frac{1}{1536\pi}. \tag{O}$$

If 3 divides M and $N = 4M/3 + 1$, then $[s, a^{4M+1}sa^{-4M-1}] = [S_0, T_{N-1}]$. Consequently there is no constant C such that $\|([S_i, T_j] - I)\Psi\| \leq CN^u e$ for all $i \neq j$, all N and all original source tuples of precision e with exact mirrors, if $u < 2$.

Proof. The statements other than (O) are Proposition 13.9, which holds for every legal choice of trees. In the collective model s is the qubit swap $\Pi((1\ 2))$ and $t = asa^{-1}$ acts at every clock as $\Pi((2\ 3))$: the base links carry $(1\ 2)$ to $(2\ 3)$ and the rotations commute with $\Pi((2\ 3))$ (proof of Proposition 13.9). The power a^{4M} returns every clock to itself, so $a^{4M+1}sa^{-4M-1} = a^{4M}ta^{-4M}$ acts at the clock c as $W_c t W_c^*$. With $h = \|W_c - I\|_{\mathrm{op}}$, $\|W_c t W_c^* - t\|_{\mathrm{op}} \leq 2h$, and since s is an involution the group commutators satisfy $\|[s, W_c t W_c^*] - [s,t]\|_{\mathrm{op}} \leq 4h$. The commutator $[s,t]$ is the qubit permutation of a three-cycle of three tensor factors, of normalized trace $\frac{1}{4}$, so $\|[s,t] - I\|_2^2 = 2 - 2\cdot\frac{1}{4} = \frac{3}{2}$. On the rows $3 \leq j \leq \lfloor M/(192\pi)\rfloor$, Lemma G.1 gives $h \leq \frac{1}{8}$, so the normalized defect at each of their clocks is at least $\sqrt{3/2} - \frac{1}{2} > \frac{1}{2}$; for $M \geq 1152\pi$ these rows number at least $M/(384\pi)$ and carry a fraction at least $1/(384\pi)$ of the clocks. On the maximally entangled state the squared residual is the average over clocks of the squared normalized defect, which proves (O). If $3 \mid M$, then $3(N-1)+1 = 4M+1$, so the word is $[S_0, T_{N-1}]$, with $N \leq 2M$; and the defect divided by $N^u e_M$ is at least $M^{2-u}/(8\pi 2^u\sqrt{1536\pi})$, unbounded for $u < 2$. □

The theorem concerns the raw worst-pair architecture, not the game demand: these models carry $\Theta(M) = \Theta(e_M^{-1/2})$ ebits, the order of the quarter itself. In them the shadow of $[S_0, T_{N-1}]$ in the plane of a and b uses only a-edges and has zero winding at every plaquette, so its signed area vanishes; a bound by signed area plus perimeter cannot give the exponent $\frac{3}{2}$ either. Every filling of $[s, a^{4M+1}sa^{-4M-1}]$ needs at least $M^2/(8\pi\sqrt{1536\pi})$ cells H4, since all other relators are exact and the operator norm of the word is at least its normalized defect; [55, Proposition 3.4] gives a stronger quadratic lower bound for these words.

At $N = 4M/3 + 1$ the undiluted models have $N^4 e_M^2$ bounded, while the window proof uses $N \asymp e^{-2/(2u+1)}$, at which $N^4 e^2 \to 0$ for $u > \frac{3}{2}$. Diluting the defect on a flag sector of weight decreasing with M reaches that scale.

Proposition G.3 (The obstruction at the scale of the window proof)**.** *Let $\frac{3}{2} < u < 2$. For sufficiently large M divisible by 3 put $p = M^{3-2u}$, choose $M' = \lceil p^{-1/4} M\rceil$, and let*

$\Psi = \sqrt{p}\,\Phi_{D_M} \oplus \sqrt{1-p}\,\Phi_{D_{M'}}$, *the two summands in orthogonal local flag sectors of both parties, with the collective model of Theorem G.2 at* M *on the first and a collective model at* M' *on the second, block diagonal, and Alice's transposes. Then* $J = -I$*, every relator other than H4 is exact, every mirror is exact, every state commutator vanishes, and the precision is at most* $\bar{e} = 8\sqrt{2}\pi\, M^{-(2u+1)/2}$*. With* $N = 4M/3 + 1$*, so that* $N \asymp \bar{e}^{-2/(2u+1)}$*,*

$$\frac{\|([S_0, T_{N-1}] - I)\Psi\|}{N^u \bar{e}} \;\geq\; \frac{M^{2-u}}{2^u\, 8\sqrt{2}\pi\sqrt{1536\pi}} \;\longrightarrow\; \infty.$$

The original state has entanglement $h(p) + p \log D_M + (1-p)\log D_{M'} = O(M') = O(\bar{e}^{-1/2})$*.*

Proof. The operators are block diagonal for the flags, so all exact relations stay exact, the mirrors of Alice's transposes are exact on each summand, and Bob's marginal is a direct sum of multiples of identities on the two blocks, with which block-diagonal operators commute. The squared H4 residual on Ψ is at most $p(8\pi/M^2)^2 + (1-p)(8\pi/M'^2)^2 \leq 2p(8\pi/M^2)^2$, because $M'^4 \geq M^4/p$; this is $\bar{e}^2$ with $p = M^{3-2u}$. The squared raw residual is at least $p/(1536\pi)$ by (O), from the first block. Dividing $\sqrt{p/(1536\pi)}$ by $N^u\bar{e} \leq (2M)^u\, 8\sqrt{2}\pi\, M^{-(2u+1)/2}$ and using $\sqrt{p} = M^{(3-2u)/2}$ gives the displayed ratio. Finally $M \asymp \bar{e}^{-2/(2u+1)}$ and, since $p^{-1/4}M \leq M' \leq p^{-1/4}M + 1 \leq 2p^{-1/4}M$, $M' \asymp p^{-1/4}M = M^{(2u+1)/4} \asymp \bar{e}^{-1/2}$; with $\log D_n = 4n + 4 + 2\log n$ (Theorem G.2) the entanglement is $O(M')$. □

So the worst-pair raw cross-copy route, whatever area bound feeds it, cannot give a lower exponent beyond $\frac{1}{5}$, the value at $u = 2$; a quadratic area bound would reach exactly that value and no more. The diluted models do not refute the quarter: their entanglement is of the quarter order. They also leave the same-state quartic table target of Proposition 15.5 open, since an averaged table defect sees the bad flag sector only with its weight.

G.1 The six-type filling construction

This proves Proposition 13.6a. The retained-relator, doubling and rewriting method is credited to [55, Appendix A and Remark A.3]; see [84, Section 4.2] for the broader filling framework. The choice of six types and the raw block order below give the stated matrix.

Proof. All areas in this proof are ultimately measured in the original five Johnson relators. No relator is added to the game.

Put

$$\sigma_i = a^i s a^{-i}, \qquad q = \sigma_{-2}, \qquad t = b^{-1}a, \qquad h_n = b^{-n}a^n.$$

Write

$$A(d) = \mathrm{Area}[s, \sigma_d].$$

Conjugation, inversion, and inversion of either commutator entry preserve area. Therefore the far-commutator problem reduces to bounding $A(d)$ for $2 \leq d \leq R$.

Six relation types for substitution Use the following ordered list of relation words:

$$\begin{aligned} F_2 &= [s, \sigma_2], \\ H &= aba^{-1}b^{-1}s^{-1}, \\ J_0 &= bsb^{-1}\sigma_1^{-1}, \\ N_2 &= [\sigma_{-2}, b], \\ N_3 &= [\sigma_{-3}, b], \\ J_1 &= b\sigma_1 b^{-1}\sigma_2^{-1}. \end{aligned}$$

Here J_0, J_1 are labels for Houghton relations, unrelated to the marked central generator J.

Their filling costs in the original presentation are at most

$$(1, 1, 1, 9, 12, 4).$$

The three auxiliary costs follow from the following fillings.

For N_2, two H-cells give

$$a^2ba^{-2} = \sigma_1 sb.$$

Replacing both occurrences in $[s, a^2ba^{-2}]$ costs four H-cells. One J_0-cell turns the resulting commutator into

$$s\sigma_1 s\sigma_1^{-1}s^{-1}\sigma_1^{-1}.$$

Three square cells turn this into $(s\sigma_1)^3$, and one braid cell fills it. Conjugation by a^{-2} proves N_2 at cost 9.

For N_3, conjugate N_2 by a^{-1}. In both occurrences substitute

$$a^{-1}ba = \sigma_{-1}^{-1}b$$

at cost two H-cells. The additional commutator $[\sigma_{-3}, \sigma_{-1}]$ costs one F_2-cell. Thus the cost is $9 + 2 + 1 = 12$.

For J_1, use $ba = s^{-1}ab$ and its inverse, then J_0:

$$b\sigma_1 b^{-1} = s^{-1}a(bsb^{-1})a^{-1}s = s^{-1}\sigma_2 s.$$

One F_2-cell finishes the transfer. The total cost is $2 + 1 + 1 = 4$.

During recursion these six types remain atomic. After recursion, every auxiliary cell is expanded into its fixed original filling. The resulting count is van Kampen area in the original presentation.

Elementary transfer fillings Vectors below count the six retained cell types in the displayed order. Let $\mathbf{f}_d, \mathbf{n}_k, \mathbf{j}_i, \mathbf{c}_m$ be filling vectors for

$$[s, \sigma_d], \quad [\sigma_{-k}, b], \quad b\sigma_i b^{-1}\sigma_{i+1}^{-1}, \quad \sigma_m(b^m s b^{-m})^{-1},$$

respectively.

Begin with the corresponding unit vectors:

$$\mathbf{f}_2 = \mathbf{e}_{F_2}, \quad \mathbf{n}_2 = \mathbf{e}_{N_2}, \quad \mathbf{n}_3 = \mathbf{e}_{N_3}, \quad \mathbf{j}_0 = \mathbf{e}_{J_0}, \quad \mathbf{j}_1 = \mathbf{e}_{J_1}.$$

The following constructions give componentwise upper bounds:

$$\mathbf{j}_i = \mathbf{j}_{i-1} + 2\mathbf{e}_H + \mathbf{f}_{i+1}, \qquad i \geq 2, \tag{G.F1}$$

$$\mathbf{c}_0 = 0, \qquad \mathbf{c}_m = \mathbf{c}_{m-1} + \mathbf{j}_{m-1}, \tag{G.F2}$$

$$\mathbf{f}_{k+j} \preceq \mathbf{f}_k + 2j\mathbf{n}_k + 2\mathbf{c}_j, \qquad k \geq 2,\ j \geq 1, \tag{G.F3}$$

$$\mathbf{n}_m = \mathbf{n}_{m-1} + \mathbf{f}_{m-1} + 2\mathbf{e}_H, \qquad m \geq 4. \tag{G.F4}$$

The recurrences follow from these substitutions:

- In (G.F1), substitute $ba = s^{-1}ab$ and its inverse, apply the a-conjugate of the preceding transfer, then commute s past σ_{i+1}.
- Equation (G.F2) conjugates the preceding transfer by b and applies J_{m-1}.

- For (G.F3), conjugate the target to $[\sigma_{-k}, \sigma_j]$, replace σ_j and its inverse by $b^j s b^{-j}$, and commute σ_{-k} through the $2j$ letters $b^{\pm 1}$. The remaining commutator is a conjugate of F_k.
- For (G.F4), conjugate N_{m-1} by a^{-1}, use $a^{-1}ba = \sigma_{-1}^{-1}b$ twice, and commute σ_{-m} with σ_{-1}.

Choose

$$\begin{aligned}\mathbf{f}_3 &= \mathbf{f}_2 + 2\mathbf{n}_2 + 2\mathbf{c}_1,\\ \mathbf{f}_4 &= \mathbf{f}_3 + 2\mathbf{n}_3 + 2\mathbf{c}_1,\\ \mathbf{f}_5 &= \mathbf{f}_3 + 4\mathbf{n}_3 + 2\mathbf{c}_2,\\ \mathbf{f}_6 &= \mathbf{f}_3 + 6\mathbf{n}_3 + 2\mathbf{c}_3.\end{aligned}$$

These choices are noncircular: $\mathbf{c}_3$ needs $\mathbf{j}_2$, which needs only $\mathbf{f}_3$.

They give

	F_2	H	J_0	N_2	N_3	J_1
$\mathbf{f}_2$	1	0	0	0	0	0
$\mathbf{f}_3$	1	0	2	2	0	0
$\mathbf{f}_4$	1	0	4	2	2	0
$\mathbf{f}_5$	1	0	4	2	4	2
$\mathbf{f}_6$	3	4	8	6	6	4
$\mathbf{n}_4$	1	2	2	2	1	0
$\mathbf{n}_5$	2	4	6	4	3	0
$\mathbf{n}_6$	3	6	10	6	7	2

The six-type doubling matrix Define a homomorphism of the free group by

$$D(a) = a^2, \qquad D(b) = b^2, \qquad D(s) = [a^2, b^2] = T.$$

The free identity

$$[a^2, b^2] = (a[a,b]a^{-1})(ab[a,b]b^{-1}a^{-1})[a,b](b[a,b]b^{-1})$$

gives, with four H-cells and two J_0-cells,

$$T = \Pi := \sigma_1\sigma_2\sigma_0\sigma_1.$$

Write

$$\mathbf{v} = 4\mathbf{e}_H + 2\mathbf{e}_{J_0}.$$

Keep this raw order, including the final positive σ_1. The six image fillings have vectors

$$\begin{aligned}D(F_2) &: 4\mathbf{v} + \mathbf{f}_2 + 4\mathbf{f}_3 + 6\mathbf{f}_4 + 4\mathbf{f}_5 + \mathbf{f}_6,\\ D(H) &: 0,\\ D(J_0) &: 2\mathbf{v} + \mathbf{j}_0 + 3\mathbf{j}_1 + 3\mathbf{j}_2 + \mathbf{j}_3,\\ D(N_2) &: 2\mathbf{v} + 2(2\mathbf{n}_3 + \mathbf{n}_2 + \mathbf{n}_4),\\ D(N_3) &: 2\mathbf{v} + 2(2\mathbf{n}_5 + \mathbf{n}_4 + \mathbf{n}_6),\\ D(J_1) &: 2\mathbf{v} + \mathbf{j}_2 + 3\mathbf{j}_3 + 3\mathbf{j}_4 + \mathbf{j}_5.\end{aligned} \tag{G.F5}$$

Here is the counting behind every row.

For $D(F_2) = [T, a^4 T a^{-4}]$, four conversions replace T by Π. Commuting its four factors past the four shifted factors requires sixteen far commutators. Their distances $2, 3, 4, 5, 6$ occur with multiplicities $1, 4, 6, 4, 1$.

The word $D(H)$ is freely trivial by the definition of T.

For $D(J_0)$, compare $b^2\Pi b^{-2}$ with $a^2\Pi a^{-2}$. A factor σ_i uses J_i and J_{i+1}. The indices $0,1,2$ occur in Π with multiplicities $1,2,1$, giving the stated sum.

For $D(N_2)$, the shifted factors have indices

$$-3,-2,-4,-3.$$

Commuting each through b^2 uses twice its N-filling. For $D(N_3)$, the indices are

$$-5,-4,-6,-5.$$

Finally, $D(J_1)$ uses two-step transfers on the indices $3,4,2,3$.

Substitution of (G.F1)–(G.F4) into (G.F5) yields the matrix

$$M=\begin{pmatrix} 18 & 20 & 64 & 34 & 34 & 12\\ 0 & 0 & 0 & 0 & 0 & 0\\ 5 & 18 & 17 & 10 & 2 & 7\\ 2 & 12 & 8 & 6 & 6 & 0\\ 16 & 40 & 52 & 32 & 28 & 4\\ 22 & 52 & 72 & 44 & 36 & 20 \end{pmatrix}. \tag{G.F6}$$

Square and braid cells appear only when the retained auxiliary relations are expanded at the end. They do not enter the recursive substitution. Preservation of these six relation types suffices; the argument does not require a claim that D preserves every original relator.

An all-depth rewriting of h_n Put

$$u_i=t^{-i}qt^i.$$

The identities

$$\sigma_{-3}=u_1,\qquad \sigma_{-4}=u_2$$

cost, respectively,

$$\mathbf{e}_{N_2},\qquad 3\mathbf{e}_{N_2}+\mathbf{e}_{N_3}.$$

Indeed,

$$t^{-1}qt=a^{-1}bqb^{-1}a$$

uses N_2 once. For the second identity, conjugate the first by a^{-1}. Since $a=bt$,

$$a^{-1}u_1a=t^{-1}b^{-1}u_1bt.$$

Commuting u_1 with b costs two replacements by σ_{-3} and one N_3-cell. Including the initial replacement gives the stated cost.

Define formal substitutions in t,q:

$$\begin{aligned} d(t)&=t^2q, & d(q)&=u_1qu_2u_1,\\ E(t)&=t^2qt^{-1}, & E(q)&=q,\\ o&=E\circ d. \end{aligned}$$

They represent $D(t),D(q)$, and their conjugates by b^{-1}.

The costs and freely reduced formal lengths are

	F_2	H	J_0	N_2	N_3	J_1	length
$d(t)$	0	1	0	0	0	0	3
$d(q)$	0	4	2	5	1	0	10
$o(t)$	0	3	0	1	0	0	7
$o(q)$	0	10	2	9	1	0	26

(G.F7)

To verify these costs, $h_2 = t^2 q$ uses one H-cell. The conversion for $D(q)$ gives

$$\sigma_{-3} q \sigma_{-4} \sigma_{-3},$$

so its cost is $\mathbf{v} + 5\mathbf{e}_{N_2} + \mathbf{e}_{N_3}$. Conjugation by b^{-1} acts on t at one H-cell and on q at one N_2-cell. The reduced word $d(q)$ contains six signed t-letters and four signed q-letters, which gives the last row. The third row follows from the two t-letters and one q-letter of $d(t)$.

Now set

$$W_1 = t, \qquad W_{2m} = d(W_m), \qquad W_{2m+1} = o(W_m)t,$$

freely reducing after substitution.

The free identities

$$h_{2m} = D(h_m), \qquad h_{2m+1} = b^{-1} D(h_m) b t$$

prove by induction that $h_n = W_n$ has a certificate using the six retained relation types for every n.

Let $\mathbf{z}_n$ be its row vector of cell counts and let $\ell_n = |W_n|$. Equations (G.F6)–(G.F7) imply

$$\mathbf{z}_n \preceq \mathbf{z}_{\lfloor n/2 \rfloor} M + \ell_{\lfloor n/2 \rfloor} \mathbf{b}, \qquad \ell_n \le 26 \ell_{\lfloor n/2 \rfloor} + 1, \tag{G.F8}$$

where one may take

$$\mathbf{b} = (0, 10, 2, 9, 1, 0).$$

The t-exponent of W_n is n, so $\ell_n \ge 1$. Hence

$$\ell_n \le 27 \ell_{\lfloor n/2 \rfloor}.$$

Solve the recurrence The characteristic polynomial is

$$\det(tI - M) = t\big(t^5 - 89t^4 + 824t^3 - 3416t^2 - 6672t + 19648\big).$$

Deleting the zero row and its corresponding column leaves a nonnegative 5×5 matrix whose square is strictly positive. Its Perron eigenvalue λ is therefore simple and strictly dominates the moduli of its other eigenvalues.

An entirely integer upper certificate is

$$w = (200, 1, 53, 24, 151, 248)^T,$$

for which

$$Mw = (15938, 0, 4197, 1886, 11984, 19720)^T \le 80w.$$

Thus $\lambda \le 80$. If Q denotes the degree-five polynomial above, then

$$Q(79) = -5063370, \qquad Q(80) = 30871488.$$

There is an eigenvalue in $(79, 80)$, so

$$79 < \lambda < 80, \qquad \|M^k\| = O(\lambda^k).$$

For $k = \lfloor \log_2 n \rfloor$, unrolling (G.F8) bounds the cell counts by a constant times

$$\sum_{j=0}^{k-1} 27^j \lambda^{k-1-j} = O(\lambda^k),$$

because $\lambda > 27$. Expansion of the auxiliary cells costs only the fixed linear functional with coefficients $(1, 1, 1, 9, 12, 4)$. Therefore

$$\mathrm{Area}(h_n W_n^{-1}) = O(n^p), \qquad p = \log_2 \lambda.$$

Both $[s, t]$ and $[s, q]$ cost one original cell: respectively a conjugate of $J_0^{\pm 1}$ and a conjugate of F_2. Replacing the two occurrences of h_n by W_n, then commuting s past its letters, gives

$$U(n) := \mathrm{Area}[h_n, s] = O(n^p). \tag{G.F9}$$

The transfer relation

$$\sigma_n = b^n s b^{-n}$$

has a filling of cost $U(n)$, since its relation word is conjugate to $[h_n, s]$. Applying (G.F3) with $k = 2, j = n$, and expanding N_2 at cost 9, gives

$$A(n + 2) \le 1 + 18n + 2U(n).$$

Together with (G.F9), this proves

$$\boxed{A^*(R) = O(R^{\log_2 \lambda})}.$$

Numerically,

$$p = 6.30646825401196\ldots.$$

□

G.2 Fixed visible holonomy and its subdivisions

Proof of Proposition 13.9a. Assume $[\rho, V] = 0$ and simultaneously diagonalize them, $\rho v_k = r_k v_k$, $V v_k = e^{i\theta_k} v_k$, $|\theta_k| \le \pi$. Set $w_j = F_A \cdots F_{j+1} v_k$, $0 \le j \le A$. Then $w_0 = e^{i\theta_k} v_k$, $w_A = v_k$, every w_j is unit, and exactly

$$w_{j-1} - w_j = F_A \cdots F_{j+1}(F_j - I) v_k.$$

Left unitarity gives the same-input norm $\|w_{j-1} - w_j\| = \|(F_j - I)v_k\|$; both sides use the same input vector. The twisted boundary is $w_A = e^{-i\theta_k} w_0$. Its Fourier frequencies are $(-\theta_k + 2\pi l)/A$, $0 \le l < A$. The least discrete-energy eigenvalue is $4\sin^2(\theta_k/(2A))$ by the magnetic-cycle computation [40, Lemma 2.9]; the scalar equal-distribution formula is also [21, Proposition 4.6(b)]. Parseval and $\sum_{j=0}^{A-1} \|w_j\|^2 = A$ give

$$\sum_{j=1}^{A} \|(F_j - I)v_k\|^2 \ge 4A \sin^2 \frac{\theta_k}{2A}.$$

Multiply by r_k, sum and divide by A. On the -1 eigenspace, either endpoint $\theta = \pm\pi$ gives the same sine square. The factors $F_j = e^{iK/A}$ attain equality; competing factors need neither commute with each other nor preserve ρ.

For an arbitrary state, the exact suffix telescope is

$$V - I = \sum_{j=1}^{A} F_A \cdots F_{j+1}(F_j - I).$$

Multiply by the original $\sqrt{\rho}$. Left unitarity, the triangle inequality and Cauchy–Schwarz give $\|(V - I)\sqrt{\rho}\|_2^2 \le A \sum_j \|(F_j - I)\sqrt{\rho}\|_2^2$, as claimed, without right prefixes or state-invariance assumptions.

For $\rho = I/d$, $V = I - 2Q$ and $K = \pi Q$, the commuting bound and its attained equality yield

$$\inf_{F_A \cdots F_1 = I - 2Q} \frac{1}{A} \sum_j \|F_j - I\|_{2,d}^2 = 4\tau(Q) \sin^2 \frac{\pi}{2A}.$$

A product of $M-1$ disjoint qubit swaps has normalized trace $2^{-(M-1)}$: its tensor permutation has $M-1$ fewer cycles than the identity. The negative spectral projection therefore has trace $(1 - 2^{-(M-1)})/2$. The bulk area is $A_B = M(M-1)/2$, so the exact optimum is

$$2(1 - 2^{-(M-1)}) \sin^2 \frac{\pi}{M(M-1)} = \Theta(M^{-4}).$$

Subdividing this fixed visible holonomy into $O(M^2)$ factors cannot improve the M^{-2} root-defect order. The infimum includes every factorization and logarithm branch. This leaves changes to the holonomy, representation, visibility or factor-count-versus-resource relation open. For general original states, relating transported based-face errors to raw game-relator errors still needs the actual mirror and state-commutator accounting. □

References

[1] Herbert Abels. An example of a finite presented solvable group. In *Homological Group Theory (Proc. Sympos., Durham, 1977)*, volume 36 of *London Math. Soc. Lecture Note Ser.*, pages 205–212. Cambridge Univ. Press, 1979. doi:10.1017/CBO9781107325449.008.

[2] Mohammad A. Alhejji and Graeme Smith. A tight uniform continuity bound for equivocation. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 2270–2274, 2020. arXiv:1909.00787, doi:10.1109/ISIT44484.2020.9174350.

[3] Rotem Arnon-Friedman and Jean-Daniel Bancal. Device-independent certification of one-shot distillable entanglement. *New J. Phys.*, 21(3):033010, 2019. arXiv:1712.09369, doi:10.1088/1367-2630/aafef6.

[4] Goulnara Arzhantseva and Pierre-Alain Cherix. Quantifying metric approximations of discrete groups. *Ann. Fac. Sci. Toulouse Math. (6)*, 33(3):831–877, 2024. arXiv:2008.12954, doi:10.5802/afst.1788.

[5] Uri Bader, Alexander Lubotzky, Roman Sauer, and Shmuel Weinberger. Stability and instability of lattices in semisimple groups. *J. Anal. Math.*, 151(1):1–23, 2023. Result numbers refer to arXiv version 2. arXiv:2303.08943, doi:10.1007/s11854-023-0329-5.

[6] Alan Baker and Gisbert Wüstholz. Logarithmic forms and group varieties. *J. Reine Angew. Math.*, 442:19–62, 1993. doi:10.1515/crll.1993.442.19.

[7] Dominique Bakry, Thierry Coulhon, Michel Ledoux, and Laurent Saloff-Coste. Sobolev inequalities in disguise. *Indiana Univ. Math. J.*, 44(4):1033–1076, 1995. doi:10.1512/iumj.1995.44.2019.

[8] Pedro Baptista, Ranyiliu Chen, Jędrzej Kaniewski, David Rasmussen Lolck, Laura Mančinska, Thor Gabelgaard Nielsen, and Simon Schmidt. A mathematical foundation for self-testing: lifting common assumptions. *Ann. Henri Poincaré*, 2025. Published online. Result numbers refer to arXiv version 1. arXiv:2310.12662, doi:10.1007/s00023-025-01642-3.

[9] Howard Barnum and Emanuel Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *J. Math. Phys.*, 43(5):2097–2106, 2002. We cite arXiv version 2 (January 2026), which corrects the proof of the main theorem for ensembles of more than one density matrix, the case containing the classical fidelity. arXiv:quant-ph/0004088, doi:10.1063/1.1459754.

[10] Salman Beigi. A new quantum data processing inequality. *J. Math. Phys.*, 54(8):082202, 2013. We cite arXiv version 5 (2023), which removes an earlier claim that optimal operators can be chosen Hermitian. arXiv:1210.1689, doi:10.1063/1.4818985.

[11] Salman Beigi. Separation of quantum, spatial quantum, and approximate quantum correlations. *Quantum*, 5:389, 2021. arXiv:2004.11103, doi:10.22331/q-2021-01-28-389.

[12] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046–2052, 1996. arXiv:quant-ph/9511030, doi:10.1103/PhysRevA.53.2046.

[13] Charles H. Bennett, David P. DiVincenzo, and John A. Smolin. Capacities of quantum erasure channels. *Phys. Rev. Lett.*, 78(16):3217–3220, 1997. arXiv:quant-ph/9701015, doi:10.1103/PhysRevLett.78.3217.

[14] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54(5):3824–3851, 1996. arXiv:quant-ph/9604024, doi:10.1103/PhysRevA.54.3824.

[15] Dave Benson, Caterina Campagnolo, Andrew Ranicki, and Carmen Rovi. Cohomology of symplectic groups and Meyer's signature theorem. *Algebr. Geom. Topol.*, 18:4069–4091, 2018. arXiv:1710.04851, doi:10.2140/agt.2018.18.4069.

[16] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M. Renes, and Renato Renner. The uncertainty principle in the presence of quantum memory. *Nature Phys.*, 6(9):659–662, 2010. arXiv:0909.0950, doi:10.1038/nphys1734.

[17] Mario Berta, Patrick J. Coles, and Stephanie Wehner. Entanglement-assisted guessing of complementary measurement outcomes. *Phys. Rev. A*, 90(6):062127, 2014. arXiv:1302.5902, doi:10.1103/PhysRevA.90.062127.

[18] Archishna Bhattacharyya, Arthur Mehta, and Yuming Zhao. On the undecidability of quantum channel capacities, 2026. Result numbers refer to the cited arXiv text. URL: https://arxiv.org/abs/2601.22471v3, arXiv:2601.22471v3.

[19] Rutvij Bhavsar, Lewis Wooltorton, and Joonwoo Bae. Composable framework for device-independent state certification. *PRX Quantum*, 6(4):040356, 2025. arXiv:2505.06833, doi:10.1103/1j5f-qjhx.

[20] Holger Boche and Gisbert Janßen. Resource cost results for one-way entanglement distillation and state merging of compound and arbitrarily varying quantum sources. *J. Math. Phys.*, 55(8):082208, 2014. arXiv:1401.6063, doi:10.1063/1.4893635.

[21] Michel Bonnefont, Sylvain Golénia, Matthias Keller, Shiping Liu, and Florentin Münch. Magnetic-Sparseness and Schrödinger Operators on Graphs. *Ann. Henri Poincaré*, 21(5):1489–1516, 2020. doi:10.1007/s00023-020-00885-6.

[22] Martin R. Bridson. The geometry of the word problem. In Martin R. Bridson and Simon M. Salamon, editors, *Invitations to Geometry and Topology*, volume 7 of *Oxford Graduate Texts in Mathematics*, pages 29–91. Oxford University Press, Oxford, 2002. doi:10.1093/oso/9780198507727.003.0002.

[23] Gergely Bunth, József Pitrik, Tamás Titkos, and Dániel Virosztek. The swap transpose on couplings translates to Petz' recovery map on quantum channels, 2025. Result numbers refer to arXiv version 2. arXiv:2512.04919.

[24] Sabine Burgdorf and Igor Klep. The truncated tracial moment problem. *J. Operator Theory*, 68(1):141–163, 2012. arXiv:1001.3679.

[25] Francesco Buscemi and Nilanjana Datta. Distilling entanglement from arbitrary resources. *J. Math. Phys.*, 51(10):102201, 2010. doi:10.1063/1.3483717.

[26] Matthew Can and Mark M. Wilde. One-shot entanglement generation and distillation via Petz recovery and binary hypothesis testing, 2026. arXiv:2610.31501.

[27] Jorge Castillejos, Samuel Evington, Aaron Tikuisis, and Stuart White. Uniform property Gamma. *Int. Math. Res. Not. IMRN*, 2022(13):9864–9908, 2022. arXiv:1912.04207, doi:10.1093/imrn/rnaa282.

[28] Rui Chao, Ben W. Reichardt, Chris Sutherland, and Thomas Vidick. Overlapping qubits. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, volume 67 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:21, 2017. arXiv:1701.01062, doi:10.4230/LIPIcs.ITCS.2017.48.

[29] Rui Chao, Ben W. Reichardt, Chris Sutherland, and Thomas Vidick. Test for a large amount of entanglement, using few measurements. *Quantum*, 2:92, 2018. arXiv:1610.00771, doi:10.22331/q-2018-09-03-92.

[30] Michael Chapman, Thomas Vidick, and Henry Yuen. Efficiently stable presentations from error-correcting codes. *Discrete Anal.*, 2026:Paper No. 1, 2026. Result numbers refer to arXiv version 2. arXiv:2311.04681.

[31] Ranyiliu Chen and Yuming Zhao. A non-robust quantum correlation self-test, 2026. Result numbers refer to arXiv version 2. arXiv:2609.25117.

[32] Richard Cleve, Peter Høyer, Ben Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of the 19th IEEE Annual Conference on Computational Complexity (CCC 2004)*, pages 236–249. IEEE, 2004. We cite arXiv version 2 (2010), which corrects a lemma of the proceedings version and gives a new proof of its Theorem 3. arXiv:quant-ph/0404076, doi:10.1109/CCC.2004.1313847.

[33] Richard Cleve, Li Liu, and William Slofstra. Perfect commuting-operator strategies for linear system games. *J. Math. Phys.*, 58(1):012202, 2017. arXiv:1606.02278, doi:10.1063/1.4973422.

[34] Richard Cleve and Rajat Mittal. Characterization of binary constraint system games. In *Automata, Languages, and Programming (ICALP 2014)*, volume 8572 of *Lecture Notes in Comput. Sci.*, pages 320–331. Springer, 2014. arXiv:1209.2729, doi:10.1007/978-3-662-43948-7_27.

[35] Cyril Cohen and Assia Mahboubi. Formal proofs in real algebraic geometry: from ordered fields to quantifier elimination. *Log. Methods Comput. Sci.*, 8(1):1:02, 2012. arXiv:1201.3731, doi:10.2168/LMCS-8(1:2)2012.

[36] Andrea Coladangelo. A two-player dimension witness based on embezzlement, and an elementary proof of the non-closure of the set of quantum correlations. *Quantum*, 4:282, 2020. arXiv:1904.02350, doi:10.22331/q-2020-06-18-282.

[37] Andrea Coladangelo. The I_{3322} Bell inequality requires infinite dimensions, 2026. Result numbers refer to arXiv version 2. arXiv:2609.06038.

[38] Andrea Coladangelo and Jalex Stark. Robust self-testing for linear constraint system games, 2017. We cite arXiv version 2 (2019). arXiv:1709.09267.

[39] Andrea Coladangelo and Jalex Stark. An inherently infinite-dimensional quantum correlation. *Nature Commun.*, 11:3335, 2020. arXiv:1804.05116, doi:10.1038/s41467-020-17077-9.

[40] Yves Colin de Verdière, Nabila Torki-Hamza, and Françoise Truc. Essential self-adjointness for combinatorial Schrödinger operators III—Magnetic fields. *Ann. Fac. Sci. Toulouse Math. (6)*, 20(3):599–611, 2011. doi:10.5802/afst.1319.

[41] Alain Connes. Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$. *Ann. of Math. (2)*, 104(1):73–115, 1976. doi:10.2307/1971057.

[42] Yves Cornulier. Sofic profile and computability of Cremona groups. *Michigan Math. J.*, 62(4):823–841, 2013. arXiv:1305.0993, doi:10.1307/mmj/1387226167.

[43] Yves Cornulier and Romain Tessera. Dehn function and asymptotic cones of Abels' group. *J. Topol.*, 6(4):982–1008, 2013. arXiv:1203.4696, doi:10.1112/jtopol/jtt019.

[44] Toby Cubitt, David Elkouss, William Matthews, Maris Ozols, David Pérez-García, and Sergii Strelchuk. Unbounded number of channel uses may be required to detect quantum capacity. *Nature Commun.*, 6(1):6739, 2015. URL: https://arxiv.org/abs/1408.5115, doi:10.1038/ncomms7739.

[45] Eric Culf. Almost perfect strategies for projection games are approximately tracial, 2026. arXiv:2603.14746.

[46] Eric Culf and Kieran Mastel. RE-completeness of entangled constraint satisfaction problems. In *Proceedings of the 66th IEEE Annual Symposium on Foundations of Computer Science (FOCS 2025)*, pages 2194–2230. IEEE, 2025. Result numbers refer to arXiv version 2, which corrects an error of version 1. arXiv:2410.21223, doi:10.1109/FOCS63196.2025.00116.

[47] Marcus De Chiffre, Lev Glebsky, Alexander Lubotzky, and Andreas Thom. Stability, cohomology vanishing, and nonapproximable groups. *Forum Math. Sigma*, 8:e18, 2020. arXiv:1711.10238, doi:10.1017/fms.2020.5.

[48] Yves de Cornulier. Finitely presented wreath products and double coset decompositions. *Geom. Dedicata*, 122(1):89–108, 2006. arXiv:math/0509090, doi:10.1007/s10711-006-9061-4.

[49] Mikael de la Salle. Spectral gap and stability for groups and non-local games. *J. Éc. polytech. Math.*, 12:1417–1444, 2025. arXiv:2204.07084, doi:10.5802/jep.314.

[50] Pierre Deligne. Extensions centrales non résiduellement finies de groupes arithmétiques. *C. R. Acad. Sci. Paris Sér. A-B*, 287(4):A203–A208, 1978.

[51] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. A*, 461(2053):207–235, 2005. arXiv:quant-ph/0306078, doi:10.1098/rspa.2004.1372.

[52] Irit Dinur, David Steurer, and Thomas Vidick. A parallel repetition theorem for entangled projection games. *Comput. Complexity*, 24(2):201–254, 2015. Result numbers refer to arXiv version 2; its Lemma 6 is Lemma 4 of version 1. arXiv:1310.4113, doi:10.1007/s00037-015-0098-3.

[53] Seth Douglas. The I3322 quantum value is attained spatially but not in finite dimension, 2026. Result numbers refer to arXiv version 2, whose mathematical content is that of version 4.1.0, doi:10.5281/zenodo.22698978. arXiv:2609.05555.

[54] Seth Douglas and Nidhal Mghirbi. Amenable groups with nearly exponential sofic profile, and quantum channels that need nearly linear memory, 2026. Result numbers refer to arXiv version 1, identical to version 1.1.2, doi:10.5281/zenodo.23111927. arXiv:2610.04271.

[55] Seth Douglas and Nidhal Mghirbi. Houghton's group H_3 has superpolynomial sofic profile, 2026. Version 1.1.2, Zenodo. Result numbers refer to that version. doi:10.5281/zenodo.23070869.

[56] Seth Douglas and Nidhal Mghirbi. The memory and purity cost of immediate delivery, 2026. Version 1.0.0, Zenodo. doi:10.5281/zenodo.23198907.

[57] Seth Douglas and Nidhal Mghirbi. The uncomputable price of precision, 2026. Version 3.0.0, Zenodo. Unrefereed preprint; result numbers refer to this version. doi:10.5281/zenodo.23736188.

[58] Souvik Dutta and Thomas Faulkner. A canonical purification for the entanglement wedge cross-section. *J. High Energy Phys.*, 2021(3):178, 2021. arXiv:1905.00577, doi:10.1007/JHEP03(2021)178.

[59] Ken Dykema, Vern I. Paulsen, and Jitendra Prakash. Non-closure of the set of quantum correlations via graphs. *Comm. Math. Phys.*, 365(3):1125–1142, 2019. arXiv:1709.05032, doi:10.1007/s00220-019-03301-1.

[60] Marco Fanizza, Larissa Kroell, Arthur Mehta, Connor Paddock, Denis Rochette, William Slofstra, and Yuming Zhao. The NPA hierarchy does not always attain the commuting operator value, 2025. Result numbers refer to arXiv version 4. arXiv:2510.04943.

[61] P. A. Fillmore. On similarity and the diagonal of a matrix. *Amer. Math. Monthly*, 76(2):167–169, 1969. doi:10.1080/00029890.1969.12000162.

[62] Tobias Fritz. On infinite-dimensional state spaces. *J. Math. Phys.*, 54(5):052107, 2013. arXiv:1202.3817, doi:10.1063/1.4807079.

[63] Honghao Fu. Constant-sized correlations are sufficient to self-test maximally entangled states with unbounded dimension. *Quantum*, 6:614, 2022. arXiv:1911.01494, doi:10.22331/q-2022-01-03-614.

[64] Jingliang Gao. Quantum union bounds for sequential projective measurements. *Phys. Rev. A*, 92(5):052331, 2015. arXiv:1410.5688, doi:10.1103/PhysRevA.92.052331.

[65] A. Gelfond. Sur le septième problème de Hilbert. *Bull. Acad. Sci. URSS, Cl. Sci. Math. Nat. (VII sér.), no. 4*, pages 623–634, 1934.

[66] W. T. Gowers and O. Hatami. Inverse and stability theorems for approximate representations of finite groups. *Sb. Math.*, 208(12):1784–1817, 2017. arXiv:1510.04085, doi:10.1070/SM8872.

[67] Samuel J. Harris. Universality of graph homomorphism games and the quantum coloring problem. *Ann. Henri Poincaré*, 25(10):4321–4356, 2024. Result numbers refer to arXiv version 2. arXiv:2305.18116, doi:10.1007/s00023-024-01422-5.

[68] Masahito Hayashi and Keiji Matsumoto. Variable length universal entanglement concentration by local operations and its application to teleportation and dense coding, 2001. arXiv:quant-ph/0109028.

[69] Chris Heunen and Sean Tull. Categories of relations as models of quantum theory. In *Proceedings of the 12th International Workshop on Quantum Physics and Logic (QPL 2015)*, volume 195 of *Electron. Proc. Theor. Comput. Sci.*, pages 247–261, 2015. doi:10.4204/EPTCS.195.18.

[70] Michael Hochman and Tom Meyerovitch. A characterization of the entropies of multidimensional shifts of finite type. *Ann. of Math. (2)*, 171(3):2011–2038, 2010. doi:10.4007/annals.2010.171.2011.

[71] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Probl. Inf. Transm.*, 9(3):177–183, 1973. Translated from Problemy Peredachi Informatsii 9(3), 3–11.

[72] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: is there a "bound" entanglement in nature? *Phys. Rev. Lett.*, 80(24):5239–5242, 1998. arXiv:quant-ph/9801069, doi:10.1103/PhysRevLett.80.5239.

[73] Raban Iten, Joseph M. Renes, and David Sutter. Pretty good measures in quantum information theory. *IEEE Trans. Inf. Theory*, 63(2):1270–1279, 2017. arXiv:1608.08229, doi:10.1109/TIT.2016.2639521.

[74] Zhengfeng Ji, Debbie Leung, and Thomas Vidick. A three-player coherent state embezzlement game. *Quantum*, 4:349, 2020. arXiv:1802.04926, doi:10.22331/q-2020-10-26-349.

[75] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP* = RE, 2020. Version 4 (2026), accepted to Ann. of Math. arXiv:2001.04383.

[76] Aimee S. A. Johnson. Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers. *Israel J. Math.*, 77(1–2):211–240, 1992. doi:10.1007/BF02808018.

[77] D. L. Johnson. Embedding some recursively presented groups. In *Groups St Andrews 1997 in Bath, Vol. II*, volume 261 of *London Math. Soc. Lecture Note Ser.*, pages 410–416. Cambridge Univ. Press, 1999. doi:10.1017/CBO9780511666148.004.

[78] Jędrzej Kaniewski, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty from effective anticommutators. *Phys. Rev. A*, 90(1):012332, 2014. arXiv:1402.5722, doi:10.1103/PhysRevA.90.012332.

[79] Irving Kaplansky. A theorem on rings of operators. *Pacific J. Math.*, 1(2):227–232, 1951. doi:10.2140/pjm.1951.1.227.

[80] Prem Nigam Kar. Robust self-testing for synchronous correlations and games, 2025. arXiv:2503.23500.

[81] Michael Keyl, Taku Matsui, Dirk Schlingemann, and Reinhard F. Werner. Entanglement, Haag-duality and type properties of infinite quantum spin chains. *Rev. Math. Phys.*, 18(9):935–970, 2006. Result numbers refer to the arXiv version. arXiv:math-ph/0604071, doi:10.1142/S0129055X0600284X.

[82] Samad Khabbazi Oskouei, Stefano Mancini, and Mark M. Wilde. Union bound for quantum information processing. *Proc. R. Soc. A*, 475(2221):20180612, 2019. arXiv:1804.08144, doi:10.1098/rspa.2018.0612.

[83] Se-Jin Kim, Vern Paulsen, and Christopher Schafhauser. A synchronous game for binary constraint systems. *J. Math. Phys.*, 59(3):032201, 2018. arXiv:1707.01016, doi:10.1063/1.4996867.

[84] Sang Rae Lee. Geometry of Houghton's groups, 2012. arXiv:1212.0257.

[85] Debbie Leung, Ben Toner, and John Watrous. Coherent state exchange in multi-prover quantum interactive proof systems. *Chicago J. Theoret. Comput. Sci.*, 2013:Article 11, 2013. arXiv:0804.4118, doi:10.4086/cjtcs.2013.011.

[86] Debbie Leung and Bingjie Wang. Characteristics of universal embezzling families. *Phys. Rev. A*, 90(4):042331, 2014. arXiv:1311.6842, doi:10.1103/PhysRevA.90.042331.

[87] Elliott H. Lieb. Convex trace functions and the Wigner–Yanase–Dyson conjecture. *Adv. Math.*, 11(3):267–288, 1973. doi:10.1016/0001-8708(73)90011-X.

[88] Göran Lindblad. Expectations and entropy inequalities for finite quantum systems. *Comm. Math. Phys.*, 39(2):111–119, 1974. doi:10.1007/BF01608390.

[89] Göran Lindblad. Completely positive maps and entropy inequalities. *Comm. Math. Phys.*, 40(2):147–151, 1975. doi:10.1007/BF01609396.

[90] Hoi-Kwong Lo and Sandu Popescu. Concentrating entanglement by local actions: beyond mean values. *Phys. Rev. A*, 63(2):022301, 2001. doi:10.1103/PhysRevA.63.022301.

[91] Alexander Lubotzky and Izhar Oppenheim. Non p-norm approximated groups. *J. Anal. Math.*, 141(1):305–321, 2020. URL: https://arxiv.org/abs/1807.06790, doi:10.1007/s11854-020-0119-2.

[92] Roger C. Lyndon and Paul E. Schupp. *Combinatorial Group Theory*, volume 89 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, 1977. Reprinted in Classics in Mathematics, 2001, doi:10.1007/978-3-642-61896-3.

[93] Anatoly I. Mal'cev. On isomorphic matrix representations of infinite groups. *Mat. Sb. (N.S.)*, 8(50)(3):405–422, 1940. In Russian.

[94] Laura Mančinska, Jitendra Prakash, and Christopher Schafhauser. Constant-sized robust self-tests for states and measurements of unbounded dimension. *Comm. Math. Phys.*, 405(9):221, 2024. arXiv:2103.01729, doi:10.1007/s00220-024-05122-3.

[95] Laura Mančinska and Simon Schmidt. Counterexamples in self-testing. *Quantum*, 7:1051, 2023. arXiv:2212.11572, doi:10.22331/q-2023-07-11-1051.

[96] Laura Mančinska and Thomas Vidick. Unbounded entanglement in nonlocal games. *Quantum Inf. Comput.*, 15(15–16):1317–1332, 2015. arXiv:1402.4145, doi:10.26421/QIC15.15-16-4.

[97] Amine Marrakchi and Mikael de la Salle. Almost synchronous correlations and Tomita–Takesaki theory. *J. Operator Theory*, 96(1):171–183, 2026. arXiv:2307.08129, doi:10.7900/jot.2025jan29.2507.

[98] Matthew McKague. Self-testing in parallel with CHSH. *Quantum*, 1:1, 2017. arXiv:1609.09584, doi:10.22331/q-2017-04-25-1.

[99] Arthur Mehta, William Slofstra, and Yuming Zhao. Positivity is undecidable in tensor products of free algebras. *J. Eur. Math. Soc. (JEMS)*, 2026. Published online. Lemma 3.9 has the same number and statement in arXiv versions 1 and 2. arXiv:2312.05617, doi:10.4171/JEMS/1786.

[100] Tom Meyerovitch. Growth-type invariants for $\mathbb{Z}^d$ subshifts of finite type and arithmetical classes of real numbers. *Invent. Math.*, 184(3):567–589, 2011. Theorem numbers refer to the cited arXiv text. URL: https://arxiv.org/abs/0902.0223, doi:10.1007/s00222-010-0296-1.

[101] Nidhal Mghirbi and Seth Douglas. An explicit nonlocal game of quantum value one that needs exponential entanglement, 2026. Version 1.0.0, Zenodo. Unrefereed preprint; result numbers refer to this version. doi:10.5281/zenodo.23697255.

[102] Amin Moharramipour, Leonardo A. Lessa, Chong Wang, Timothy H. Hsieh, and Subhayan Sahu. Symmetry-enforced entanglement in maximally mixed states. *PRX Quantum*, 5(4):040336, 2024. arXiv:2406.08542, doi:10.1103/PRXQuantum.5.040336.

[103] Tobias Moroder, Jean-Daniel Bancal, Yeong-Cherng Liang, Martin Hofmann, and Otfried Gühne. Device-independent entanglement quantification and related applications. *Phys. Rev. Lett.*, 111(3):030501, 2013. doi:10.1103/PhysRevLett.111.030501.

[104] Tobias Moroder, Marcos Curty, and Norbert Lütkenhaus. One-way quantum key distribution: Simple upper bound on the secret key rate. *Phys. Rev. A*, 74(5):052301, 2006. doi:10.1103/PhysRevA.74.052301.

[105] Hamoon Mousavi, Seyed Sajjad Nezhadi, and Henry Yuen. Nonlocal games, compression theorems, and the arithmetical hierarchy. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2022)*, pages 1–11. ACM, 2022. Result numbers refer to arXiv version 2. arXiv:2110.04651, doi:10.1145/3519935.3519949.

[106] Magdalena Musat and Mikael Rørdam. Non-closure of quantum correlation matrices and factorizable channels that require infinite dimensional ancilla. *Comm. Math. Phys.*, 375(3):1761–1776, 2020. With an appendix by Narutaka Ozawa. arXiv:1806.10242, doi:10.1007/s00220-019-03449-w.

[107] Anand Natarajan and Thomas Vidick. A quantum linearity test for robustly verifying entanglement. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC 2017)*, pages 1003–1015, 2017. arXiv:1610.03574, doi:10.1145/3055399.3055468.

[108] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.*, 10(7):073013, 2008. arXiv:0803.4290, doi:10.1088/1367-2630/10/7/073013.

[109] Michael A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83(2):436–439, 1999. arXiv:quant-ph/9811053, doi:10.1103/PhysRevLett.83.436.

[110] Marcin L Nowakowski and Pawel Horodecki. A simple test for quantum channel capacity. *J. Phys. A: Math. Theor.*, 42(13):135306, 2009. Result numbers refer to the cited arXiv version 2. URL: https://arxiv.org/abs/quant-ph/0503070v2, doi:10.1088/1751-8113/42/13/135306.

[111] Dimiter Ostrev and Thomas Vidick. Entanglement of approximate quantum strategies in XOR games. *Quantum Inf. Comput.*, 18(7–8):617–631, 2018. arXiv:1609.01652, doi:10.26421/QIC18.7-8-6.

[112] Connor Paddock. Rounding near-optimal quantum strategies for nonlocal games to strategies using a maximally entangled state. *J. Math. Phys.*, 66(12):122204, 2025. Result numbers refer to arXiv version 5. arXiv:2203.02525, doi:10.1063/5.0209725.

[113] Connor Paddock and Simon Schmidt. A commuting operator self-test for exact entanglement embezzlement, 2026. Result numbers refer to arXiv version 1. arXiv:2609.38083.

[114] Connor Paddock and William Slofstra. Satisfiability problems and algebras of Boolean constraint system games. *Illinois J. Math.*, 69(1):81–107, 2025. Result numbers refer to arXiv version 2. arXiv:2310.07901, doi:10.1215/00192082-11831196.

[115] Connor Paddock, William Slofstra, Yuming Zhao, and Yangchen Zhou. An operator-algebraic formulation of self-testing. *Ann. Henri Poincaré*, 25(10):4283–4319, 2024. Result numbers refer to arXiv version 2. arXiv:2301.11291, doi:10.1007/s00023-023-01378-y.

[116] Károly F. Pál and Tamás Vértesi. Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems. *Phys. Rev. A*, 82(2):022116, 2010. arXiv:1006.3032, doi:10.1103/PhysRevA.82.022116.

[117] Arthur J. Parzygnat and James Fullwood. From time-reversal symmetry to quantum Bayes' rules. *PRX Quantum*, 4(2):020334, 2023. arXiv:2212.08088, doi:10.1103/PRXQuantum.4.020334.

[118] Vern I. Paulsen, Simone Severini, Daniel Stahlke, Ivan G. Todorov, and Andreas Winter. Estimating quantum chromatic numbers. *J. Funct. Anal.*, 270(6):2188–2222, 2016. arXiv:1407.6918, doi:10.1016/j.jfa.2016.01.010.

[119] Jef Pauwels. The quantum supremum of the I_{3322} Bell inequality is not attained in finite dimension, 2026. Result numbers refer to arXiv version 1. arXiv:2608.29734.

[120] Dénes Petz. Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Comm. Math. Phys.*, 105(1):123–131, 1986. doi:10.1007/BF01212345.

[121] Stefano Pironio, Miguel Navascués, and Antonio Acín. Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM J. Optim.*, 20(5):2157–2180, 2010. arXiv:0903.4368, doi:10.1137/090760155.

[122] Oded Regev and Thomas Vidick. Quantum XOR games. *ACM Trans. Comput. Theory*, 7(4):Article 15, 2015. arXiv:1207.4939, doi:10.1145/2799560.

[123] Joseph M. Renes. Better bounds on optimal measurement and entanglement recovery, with applications to uncertainty and monogamy relations. *Phys. Rev. A*, 96(4):042328, 2017. arXiv:1707.01114, doi:10.1103/PhysRevA.96.042328.

[124] Jacques Riguet. Relations binaires, fermetures, correspondances de Galois. *Bull. Soc. Math. France*, 76:114–155, 1948. doi:10.24033/bsmf.1401.

[125] Daniel J. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergodic Theory Dynam. Systems*, 10(2):395–406, 1990. doi:10.1017/S0143385700005629.

[126] Michel Rumin. Spectral density and Sobolev inequalities for pure and mixed states. *Geom. Funct. Anal.*, 20(3):817–844, 2010. arXiv:0906.2902, doi:10.1007/s00039-010-0075-6.

[127] Lukas Schmitt and David Sutter. Distillable entanglement is continuous, 2026. Result numbers refer to the cited arXiv text. URL: https://arxiv.org/abs/2610.35722v1, arXiv:2610.35722v1.

[128] Theodor Schneider. Transzendenzuntersuchungen periodischer Funktionen. I. Transzendenz von Potenzen. *J. Reine Angew. Math.*, 172:65–69, 1935. doi:10.1515/crll.1935.172.65.

[129] A. Seidenberg. A new decision method for elementary algebra. *Ann. of Math. (2)*, 60(2):365–374, 1954. doi:10.2307/1969640.

[130] J. R. Shoenfield. On degrees of unsolvability. *Ann. of Math. (2)*, 69(3):644–653, 1959. doi:10.2307/1970028.

[131] William Slofstra. A group with at least subexponential hyperlinear profile, 2018. arXiv:1806.05267.

[132] William Slofstra. The set of quantum correlations is not closed. *Forum Math. Pi*, 7:e1, 2019. Result numbers refer to arXiv version 2. arXiv:1703.08618, doi:10.1017/fmp.2018.3.

[133] William Slofstra. Tsirelson's problem and an embedding theorem for groups arising from non-local games. *J. Amer. Math. Soc.*, 33(1):1–56, 2020. arXiv:1606.03140, doi:10.1090/jams/929.

[134] William Slofstra and Thomas Vidick. Entanglement in non-local games and the hyperlinear profile of groups. *Ann. Henri Poincaré*, 19(10):2979–3005, 2018. Result numbers refer to arXiv version 2. arXiv:1711.10676, doi:10.1007/s00023-018-0718-y.

[135] Ryuji Takagi, Kaito Watanabe, Takaya Matsuura, Hayato Arai, and Masahito Hayashi. Universal distillation of quantum entanglement, 2026. Version 1. arXiv:2609.40215.

[136] Elmar Thoma. Die unzerlegbaren, positiv-definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe. *Math. Z.*, 85(1):40–61, 1964. doi:10.1007/BF01114877.

[137] Aaron Tikuisis, Stuart White, and Wilhelm Winter. Quasidiagonality of nuclear C*-algebras. *Ann. of Math. (2)*, 185(1):229–284, 2017. Result numbers refer to arXiv version 2. arXiv:1509.08318, doi:10.4007/annals.2017.185.1.4.

[138] Salvatore Tirone, Francesco Anna Mele, Vittorio Giovannetti, and Ludovico Lami. Universal entanglement distillation, 2026. Version 1. arXiv:2609.40080.

[139] Armin Uhlmann. The "transition probability" in the state space of a *-algebra. *Rep. Math. Phys.*, 9(2):273–279, 1976. doi:10.1016/0034-4877(76)90060-4.

[140] Armin Uhlmann. Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory. *Comm. Math. Phys.*, 54(1):21–32, 1977. doi:10.1007/BF01609834.

[141] Wim van Dam and Patrick Hayden. Universal entanglement transformations without communication. *Phys. Rev. A*, 67(6):060302, 2003. arXiv:quant-ph/0201041, doi:10.1103/PhysRevA.67.060302.

[142] Lauritz van Luijk, Alexander Stottmeister, Reinhard F. Werner, and Henrik Wilming. Pure state entanglement and von Neumann algebras. *Comm. Math. Phys.*, 406(12):296, 2025. Result numbers refer to arXiv version 2. arXiv:2409.17739, doi:10.1007/s00220-025-05465-5.

[143] Matthijs Vernooij and Yuming Zhao. Lifting the maximally-entangledness assumption in robust self-testing for synchronous games. *Quantum*, 10:2222, 2026. Result numbers refer to arXiv version 2. arXiv:2505.05994, doi:10.22331/q-2026-10-01-2222.

[144] Guifré Vidal. Entanglement of pure states for a single copy. *Phys. Rev. Lett.*, 83(5):1046–1049, 1999. doi:10.1103/PhysRevLett.83.1046.

[145] Thomas Vidick. Almost synchronous quantum correlations. *J. Math. Phys.*, 63(2):022201, 2022. Result numbers refer to arXiv version 3. arXiv:2103.02468, doi:10.1063/5.0056512.

[146] Cédric Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2009. doi:10.1007/978-3-540-71050-9.

[147] Andreas Winter. Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints. *Comm. Math. Phys.*, 347(1):291–313, 2016. Result numbers refer to the cited arXiv version 2. URL: https://arxiv.org/abs/1507.07775v2, doi:10.1007/s00220-016-2609-8.

[148] Michael M. Wolf, Toby S. Cubitt, and David Perez-Garcia. Are problems in quantum information theory (un)decidable?, 2011. Result numbers refer to the cited arXiv text. URL: https://arxiv.org/abs/1111.5425, arXiv:1111.5425.

[149] Yuming Zhao. Robust self-testing for nonlocal games with robust game algebras, 2024. arXiv:2411.03259.