

Mass Gap of Four-Dimensional Lattice Yang–Mills Theory for Every Compact Simple Group

Deep Bhattacharjee^{*},^{1,2,*} Ushashi Bhattacharya^{3,†} and James Montgomery Lockwood^{4,‡}

¹Formerly, *Electro-Gravitational Space Propulsion Laboratory (EGSPL), Bhubaneswar, Odisha 751030, India*

²Formerly, *University of Calcutta (Heramba Chandra College),
23/49 Gariahat Road, Golpark, Kolkata 700029, West Bengal, India*

³Formerly, *National Taiwan University, Taipei 106319, Taiwan*

⁴*Independent Researcher, San Antonio, USA*

(Dated: 30 October 2026)

We prove that four-dimensional lattice Yang–Mills theory with any compact simple gauge group has a unique infinite-volume state, exponential clustering and an explicit transfer-matrix gap at couplings past the curvature limit of the ungauged measure, up to $\beta_W = 0.345$ for $SU(2)$ and $\beta_W = 0.76$ for $SU(3)$, in a gauge that fixes one link of every plaquette. We prove that a Dobrushin–Shlosman estimate in one cube of fixed physical size, or any exponential decay of renormalised correlations in physical units, gives a mass gap for every continuum limit. The cube estimate is open along the continuum tuning; for $SU(N)$ it needs cubes of side at least of order $(\beta_W / \log \beta_W)^{1/4}$. For block averages, a cut-off of Balaban’s average composes across all scales on fields of small curvature for every compact group, and under the $SU(2)$ Wilson measure the large fields form clusters of vanishing physical size.

CONTENTS

Declarations

I. Introduction	1	A. The line field	58
II. The line field	6	B. The comparison lemma	58
III. Gauge fixing and the free-boundary state	7	C. The plaquette blocks	59
IV. $SU(2)$ up to one third	8	D. The blocks of Lemma 20	60
V. $SU(2)$ past one third: plaquette blocks	9	E. The certified computation	60
VI. Curvature in the gauge	11	F. The transfer matrix	61
VII. $SU(3)$ up to $3/4$	14	G. Spectral measures and Laplace transforms	61
VIII. $SU(3)$ with plaquette blocks	18	H. The character expansion	62
IX. Where the methods stop	19	I. The single-link law of $SU(2)$	62
X. The continuum step at one physical scale	21	J. Computations for the single-link law	66
XI. The open inputs	34	K. Bessel functions	67
XII. Block averages	44		
Acknowledgements	57		
Data and code availability	58		

I. INTRODUCTION

Wilson’s lattice gauge theory [77] assigns an element Q_e of a compact group $G \subset U(N)$ to every link e of $\mathbb{Z}^4$ and weighs a configuration by $\exp(N\beta \sum_p \text{Re Tr } Q_p)$, where Q_p is the ordered product of the four link variables around the plaquette p .¹

^{*} itsdeep@live.com; ^{*} Corresponding author. Secondary address: d.bhattacharjee@erl-forschung.de; GitHub contact: egspl@proton.me, github.com/creelie.

[†] rai.bhattacharya.in@gmail.com

[‡] spectralityworkinggroup@gmail.com; Secondary address: Blaze61029doqq@proton.me.
2020 *Mathematics Subject Classification*. Primary 81T25, 82B20; Secondary 60K35, 81T13, 22E30, 58J50, 65G30. *Keywords*. Lattice gauge theory; gauge fixing; Dobrushin uniqueness; mass gap. Source, checks and Lean certificate: github.com/creelie/4Y-M.

¹ This is the normalisation of Shen, Zhu and Zhu [69] and of our earlier paper [14]. The conventional inverse coupling is $\beta_W = N^2\beta$, so the weight is $\exp(\frac{\beta_W}{N} \sum_p \text{Re Tr } Q_p)$; for $SU(2)$, identified with the unit quaternions, $\frac{1}{2} \text{Tr } Q_p = \text{Re } Q_p$ and $\beta_W = 4\beta$. Standard accounts of the lattice theory are the books of Montvay and Münster [61] and of Gittering and Lang [37]; the Hamiltonian version is due to Kogut and Susskind [51].

At small $|\beta|$ the theory is under control: it has a unique infinite-volume state, correlations decay exponentially, and the transfer matrix has a gap [64, 68].² The question is how far into the coupling range such statements can be proved with explicit constants, since the problem of the continuum limit [47] begins where strong coupling ends and Bałaban's weak-coupling analysis [5–7] takes over.

Two methods give explicit ranges. The Bakry–Émery method [3, 4] of Shen, Zhu and Zhu [69] bounds the Hessian of the action; in [14] we showed that the sharp bound is $4d = 16$ and that no curvature argument for the ungauged measure goes beyond $|\beta| < 1/32$. The Dobrushin method [29, 30]³ bounds the influence of each link on the conditional law of another; for $SU(2)$ the single-link conditional law is a von Mises–Fisher law on the three-sphere, the sharp Kantorovich modulus⁴ of that family is one quarter [14] (Theorem 125 in Appendix I), and every link has eighteen partners, which gives $\beta_W < 2/9$ and nothing beyond, because the coefficients are attained.

Both limits come from the same count. A link of $\mathbb{Z}^4$ lies in six plaquettes and shares each with three other links. Gauge invariance makes part of this redundant: in a box, the variables on a forest of links can be set to the identity without changing any gauge-invariant expectation [26].⁵ The forest we use is a field of straight lines with one line through every site, chosen so that every plaquette contains exactly one fixed link. Then every free link has exactly twelve free partners, two in each of its plaquettes (Lemma 9), and both methods improve.

² Strong-coupling results of this kind go back to Osterwalder and Seiler [64], who proved exponential clustering and a gap by a convergent expansion, and to Seiler's monograph [68]; the general framework for lattice gauge fields and their continuum limits is that of Brydges, Fröhlich and Seiler [21]. More recent work treats the strong-coupling regime by probabilistic methods: the survey of Chatterjee [24], his large- N solution of $SO(N)$ lattice gauge theory at strong coupling [25] and its planar extension by Basu and Ganguly [10], the Langevin dynamics of Shen, Zhu and Zhu [69], and the area law in an enlarged region by Cao, Nissim and Sheffield [23]. For finite gauge groups Adhikari and Cao proved correlation decay at weak coupling [1]. In two dimensions the continuum Yang–Mills measure has been constructed by Gross, King and Sengupta [40], Driver [32] and Lévy [52]; in four dimensions Bałaban's renormalisation group, as expounded by Dimock [28], gives ultraviolet stability in finite volume, and Magnen, Rivasseau and Sénéor constructed the theory with an infrared cutoff [57]. None of these results gives a mass gap in physical units.

³ Dobrushin's theorem: if the conditional law of each variable given all others moves, in Kantorovich distance, by at most C_{xy} times the distance by which the variable at y moves, and $\sup_x \sum_y C_{xy} < 1$, then the Gibbs measure is unique and correlations decay at the rate given by the Neumann series of the matrix C . Dobrushin and Shlosman [31] replaced single variables by finite boxes, which is the form used in Sec. X.

⁴ The Kantorovich distance between two probability measures on a metric space is the least expected distance over their couplings, equivalently the largest difference of their integrals over 1-Lipschitz functions [75]. The modulus of a family of laws indexed by a metric space is the best Lipschitz constant of the map from the index to the law in that distance.

⁵ Creutz used a maximal tree to remove the gauge freedom completely; the temporal gauge, with all links in one direction fixed, is the special case of a family of straight lines. In a maximal tree almost every link is fixed in some region and none in others, so the number of partners of a free link is not uniform, and the temporal gauge leaves fourteen partners per link. The point here is a pattern that fixes exactly one link in every plaquette.

Theorem 1 ($SU(2)$). *Let $G = SU(2)$ and $0 \leq \beta_W < \frac{1}{3}$. For every gauge-invariant local observable O the expectations $W_\Lambda(O)$ in boxes with free boundary converge, along every sequence of boxes that exhausts $\mathbb{Z}^4$, to a limit $\omega(O)$ that does not depend on the sequence. The state ω is invariant under translations and the hypercubic group and reflection positive;⁶ truncated correlations of Lipschitz observables decay at every rate below $-\log(3\beta_W)$; and the transfer matrix has spectral gap*

$$m \geq -\log(3\beta_W) > 0.$$

For $-\frac{1}{3} < \beta_W < 0$ all statements except reflection positivity and the gap hold.

The step behind Theorem 1 is short once the line field is in place. The number $\frac{1}{3}$ is also the end of the road for this kind of argument.

Theorem 2 (Ceiling). *Let F' be any set of links of $\mathbb{Z}^4$ without cycles, and give two distinct links outside F' that share a plaquette the coefficient $|\beta|$, the largest single-link Kantorovich coefficient of $SU(2)$. Then the weighted Dobrushin condition with these coefficients fails for $\beta_W \geq \frac{1}{3}$. For the line field the coefficients are attained, so no single-link Kantorovich argument in that gauge reaches $\beta_W = \frac{1}{3}$.*

The ceiling concerns coefficients between single links. Take instead as a block the three free links of one plaquette. The free links of the line field split into such triples with period two (Lemma 12), the coupling inside the block plaquette is no longer paid for, and each link keeps ten outside partners. The price is the sensitivity of a law on three spheres instead of one, which we bound analytically by integrating out the middle link of the block.

Theorem 3 ($SU(2)$ with plaquette blocks). *Let $G = SU(2)$ and $0 \leq \beta_W \leq 0.345$. The number $\alpha(\beta_W)$ of (4) is increasing, $\alpha(\frac{1}{3}) \leq 0.9561$ and $\alpha(0.345) \leq 0.9973$. The conclusions of Theorem 1 hold with decay at every rate below $-\frac{1}{2} \log \alpha(\beta_W)$ and a transfer-matrix gap*

$$m \geq -\frac{1}{2} \log \alpha(\beta_W) > 0;$$

at $\beta_W = \frac{1}{3}$ this gives $m \geq 0.0224$. For $-0.345 \leq \beta_W < 0$ all statements except reflection positivity and the gap hold.

At small coupling the rate of Theorem 1 is the better one, so both bounds stand side by side; Theorem 3 matters past $\beta_W = \frac{1}{3}$, where the gap of Theorem 1 has closed.

The curvature method gains as well, and for every N . Let $\text{Hess } S_0$ be the Hessian of $\sum_p \text{Re Tr } Q_p$ on G^E and F the set of fixed links.

⁶ That is, $\omega(\vartheta(\bar{F})F) \geq 0$ for every F depending on the links on one side of a lattice hyperplane, ϑ being the reflection in it. It gives the Hilbert space and the transfer matrix of Osterwalder and Schrader [62, 63] and of Osterwalder and Seiler [64], reconstructed in Appendix F.

Theorem 4 (Curvature in the gauge). *For every N , every field Q and every tangent vector $v = (X_e Q_e)$ with $X_e = 0$ for $e \in F$,*

$$|\text{Hess} S_0(v, v)| \leq \frac{29}{2} |v|^2,$$

against the sharp constant 16 without gauge fixing. Consequently, for $G = \text{SU}(N)$ and $|\beta| < \frac{1}{29}$, or $G = \text{SO}(N)$, $N \geq 3$, and $|\beta| < \frac{N-2}{58N}$, the conclusions of Theorem 1 hold with the rate and gap of Theorem 5. For $\text{SU}(3)$ this is $\beta_W < \frac{9}{29} = 0.310$.

The constant $\frac{29}{2}$ is a statement about $\text{U}(N)$, so it holds on every closed subgroup, and the curvature of the group does the rest. This gives the theorem for every compact simple gauge group, the exceptional groups included, with an explicit gap. Let $\rho : G \rightarrow \text{U}(N)$ be a faithful unitary representation of a compact connected Lie group G , weigh a configuration by $\exp(N\beta \sum_p \text{Re Tr } \rho(Q_p))$, and let κ_G be the least Ricci⁷ curvature of G for the bi-invariant metric $\langle X, Y \rangle = -\text{Tr}(\rho(X)\rho(Y))$; we identify G with $\rho(G) \subset \text{U}(N)$.

Theorem 5 (Every compact simple group). *Suppose that the Lie algebra of G is semisimple, so that $\kappa_G > 0$, and that*

$$N|\beta| < \frac{29}{2} \kappa_G.$$

Then the conclusions of Theorem 1 hold, for $\beta < 0$ all of them except reflection positivity and the gap, with decay at every rate below

$$m_G(\beta) = \text{arccosh}\left(1 + \frac{1}{12} \left(\frac{\kappa_G}{N|\beta|} - \frac{29}{2} \right)\right) > 0$$

and a transfer-matrix gap $m \geq m_G(\beta)$. For a simple G , $\kappa_G = h^\vee / (2\ell_\rho)$, where $h^\vee$ is the dual Coxeter number and ℓ_ρ the Dynkin index of ρ ; Table 1 lists the values.

For $\text{SU}(N)$, $\kappa_G = \frac{N}{2}$, and for $\text{SO}(N)$, $\kappa_G = \frac{N-2}{4}$, so Theorem 5 contains the ranges of Theorem 4 and supplies their rate.⁸ For small $|\beta|$ the existence of a gap is classical for every compact group, with no explicit range [64, Thm. 3.6]; what is new is an explicit range for every group, past the range $N|\beta| < \kappa_G/16$ that the Hessian bound 16 gives without gauge fixing, and an explicit gap. As $\beta \rightarrow 0$ the gap grows like $\log \frac{\kappa_G}{6N|\beta|}$.

No Bakry–Émery argument for the ungauged measure can pass $|\beta| = \frac{1}{32}$ [14, Thm. C]; the gauge fixing changes the

TABLE I. Least Ricci curvature $\kappa_G = h^\vee / (2\ell_\rho)$ of the compact simple groups in a faithful representation ρ of least dimension N , and the range of Theorem 5 for the conventional inverse coupling $\beta_W = N^2\beta$.

Group	ρ	$h^\vee$	ℓ_ρ	κ_G	$\beta_W <$
$\text{SU}(N)$, $N \geq 2$	$\mathbb{C}^N$	N	1	$\frac{N}{2}$	$\frac{N^2}{29}$
$\text{SO}(N)$, $N \geq 5$	$\mathbb{R}^N$	$N-2$	2	$\frac{N-2}{4}$	$\frac{N(N-2)}{58}$
$\text{SO}(3)$	$\mathbb{R}^3$	2	4	$\frac{1}{4}$	$\frac{3}{58}$
$\text{Sp}(n)$, $n \geq 2$	$\mathbb{C}^{2n}$	$n+1$	1	$\frac{n+1}{2}$	$\frac{2n(n+1)}{29}$
G_2	7	4	2	1	$\frac{14}{29}$
F_4	26	9	6	$\frac{3}{4}$	$\frac{39}{29}$
E_6	27	12	6	1	$\frac{54}{29}$
E_7	56	18	12	$\frac{3}{4}$	$\frac{84}{29}$
E_8	248	30	60	$\frac{1}{4}$	$\frac{124}{29}$

measure, not the metric, which is why it can. For $\text{SU}(2)$ the curvature range $\beta_W < \frac{4}{29}$ lies inside that of Theorem 1, and no curvature argument for the gauge-fixed $\text{SU}(2)$ measure passes $\beta_W = \frac{1}{3}$: when one active link e is -1 and all others are 1, the six plaquettes through e have variable -1 , and along a geodesic through Q_e the Hessian of $-\frac{\beta_W}{2} \sum_p \text{Re Tr } Q_p$ is $-6\beta_W$ times the metric of the unit sphere S^3 , whose Ricci curvature is 2, so the Bakry–Émery tensor is negative there once $\beta_W > \frac{1}{3}$. For $\text{SU}(3)$ the Dobrushin route goes much further.

Theorem 6 ($\text{SU}(3)$). *Let $G = \text{SU}(3)$ and $0 \leq \beta_W \leq \frac{3}{4}$. Then the conclusions of Theorem 1 hold, with decay at every rate below $-\log(1.33319\beta_W)$ and a transfer-matrix gap*

$$m \geq -\log(1.33319\beta_W) > 0.$$

The proof of Theorem 6 bounds the Kantorovich modulus of the one-link measure $\nu_b(dU) \propto e^{\text{Re Tr}(U^b)} dU$ on $\text{SU}(3)$ uniformly over staple sums b . Its analytic part is a lower bound for the spectral gap of ν_b through Bochner’s formula with a cross term (Lemma 29), a corrector chosen box by box among polynomials of degree two in the matrix entries (Lemma 30), the reduction of b to diagonal matrices (Lemma 31), the character expansion of the moments of ν_b (Lemma 32) and an interpolation inequality on boxes of staple sums (Lemma 33). The remaining step evaluates the resulting bound on 2341 boxes with rigorous error bounds; the theorem rests on that computation, which is described in full in Sec. VII G and Appendix E.

The plaquette blocks of Theorem 3 work for $\text{SU}(3)$ as well. They use the same modulus of the one-link measure, but only through the three links of a block, so they need $\Lambda(2\beta_W)$ small and not $2r\Lambda(r) < 1$: near the end of Theorem 6 they give a much larger gap, and they go past it.

Theorem 7 ($\text{SU}(3)$ with plaquette blocks). *Let $G = \text{SU}(3)$ and $0 \leq \beta_W \leq \frac{19}{25} = 0.76$. Then $\alpha_3(\beta_W) < 1$ for the number of (15), with $\alpha_3(0.72) \leq 0.9182$, $\alpha_3(\frac{3}{4}) \leq 0.9653$ and $\alpha_3(0.76) \leq 0.99950$. The conclusions of Theorem 1 hold with decay at every rate below $-\frac{1}{2} \log \alpha_3(\beta_W)$ and a transfer-matrix gap*

$$m \geq -\frac{1}{2} \log \alpha_3(\beta_W) > 0;$$

⁷ For a bi-invariant metric on a compact Lie group the Ricci curvature is $-\frac{1}{4} \text{Tr}(\text{ad } X)^2$ in the direction of X . It is positive for every $X \neq 0$ exactly when the Lie algebra is semisimple, and κ_G is its minimum over unit vectors.

⁸ The dual Coxeter number is $h^\vee = N$ for $\text{SU}(N)$, $N-2$ for $\text{SO}(N)$, $n+1$ for $\text{Sp}(n)$, 4, 9, 12, 18, 30 for G_2, F_4, E_6, E_7, E_8 , and the Dynkin index of a representation is normalised to one on the defining representation of $\text{SU}(N)$ [49, Ch. 6]. The identity $\kappa_G = h^\vee / (2\ell_\rho)$ is the statement that the Ricci tensor of a compact simple group with a bi-invariant metric is one quarter of the negative Killing form, $\text{Ric}(X, X) = -\frac{1}{4} B(X, X)$ [60], together with $B(X, X) = 2h^\vee \text{Tr}(\rho(X)^2) / \ell_\rho$. Lower Ricci bounds of the same kind, for special holonomy and for Riemannian submersions, are studied in [18, 38].

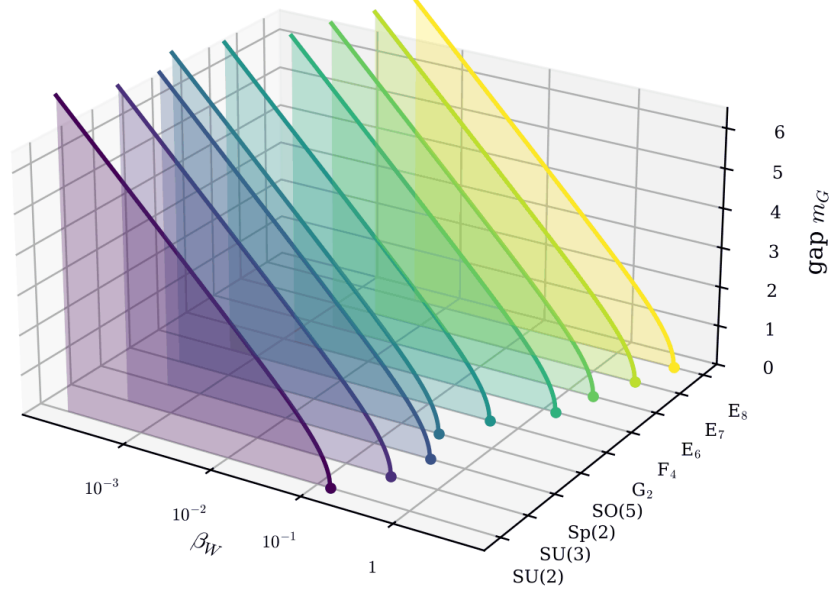


FIG. 1. The gap $m_G(\beta)$ of Theorem 5 as a function of $\beta_W = N^2\beta$ on a logarithmic scale, for nine groups in the representations of Table I. Each curve ends at the edge $\beta_W = 2N\kappa_G/29$ of the range, marked by a dot, where the gap tends to zero; as $\beta_W \rightarrow 0$ it grows like $\log(\kappa_G/(6N|\beta|))$. The curves differ only by the factor $N\kappa_G$, since m_G depends on $N|\beta|/\kappa_G$ alone.

TABLE II. Explicit ranges of uniqueness, decay and a transfer-matrix gap for four-dimensional lattice Yang–Mills theory. The first three rows hold for the ungauged measure; the last three are the results of this paper.

Method	SU(2), β_W	SU(3), β_W
Curvature, Shen–Zhu–Zhu [69]	$< \frac{1}{12}$	$< \frac{3}{16}$
Curvature, sharp [14]	$< \frac{1}{8}$	$< \frac{9}{32}$
Dobrushin, sharp quarter [14]	$< \frac{2}{9}$	
Curvature, forest gauge	$< \frac{4}{29}$	$< \frac{9}{29}$
Dobrushin, forest gauge	$< \frac{1}{3}$	$\leq \frac{3}{4}$
Plaquette blocks, forest gauge	≤ 0.345	≤ 0.76

at $\beta_W = \frac{3}{4}$ this gives $m \geq 0.0177$, against $1.1 \cdot 10^{-4}$ from Theorem 6, and at $\beta_W = 0.76$, where Theorem 6 no longer applies, $m \geq 2.5 \cdot 10^{-4}$.

All seven theorems are statements at fixed lattice spacing. They do not touch the continuum limit, the crossover to Balaban’s regime, or a gap in physical units, which together are the Clay problem; Sec. IX says where each method stops and what remains of the four open problems of [14]. The obstruction proved in [15], that a lattice gap uniform in lattice units forces trivial continuum two-point functions, applies to all of them.

The continuum step. Sec. X goes past that obstruction in one respect, through three implications, each proved.

(I1) A Dobrushin–Shlosman condition on one cube of fixed

physical size, uniform along the continuum tuning, gives a mass gap in physical units for every continuum limit (Theorems 43 and 45), and so does exponential decay with constants of the renormalised correlations of a spanning set of observables (Theorem 46); this holds also for observables smoothed by the spatial gradient flow (Proposition 47).

(I2) The flowed observables, renormalised by one condition at a single physical separation, have reflection positive continuum limits along subsequences, and such a limit is nontrivial exactly when one ratio of lattice two-point functions stays away from zero (Theorem 48); that ratio is bounded above and below by ratios of correlations of one slice field at fixed physical separations (Proposition 51).

(I3) Tree bounds on the truncated correlations give all the higher Schwinger functions as reflection positive distributions, imply the ratio bound, and carry the conditional mass gap to every polynomial in the smeared fields (Theorem 56); for a smooth packet they give the ratio with an explicit rate and a reference vector of finite energy (Proposition 60).

What is open. The inputs of (I1)–(I3), namely the cube condition along the tuning, the slice ratio in four dimensions and the tree bounds, are not proved, and so the nontrivial existence of the continuum limit and the gap in physical units remain open. Figure 3 shows how the proved statements, the implications and the open inputs fit together.

Sec. XI measures the open inputs. The cube condition holds at strong coupling, fails at every fixed cube size at weak cou-

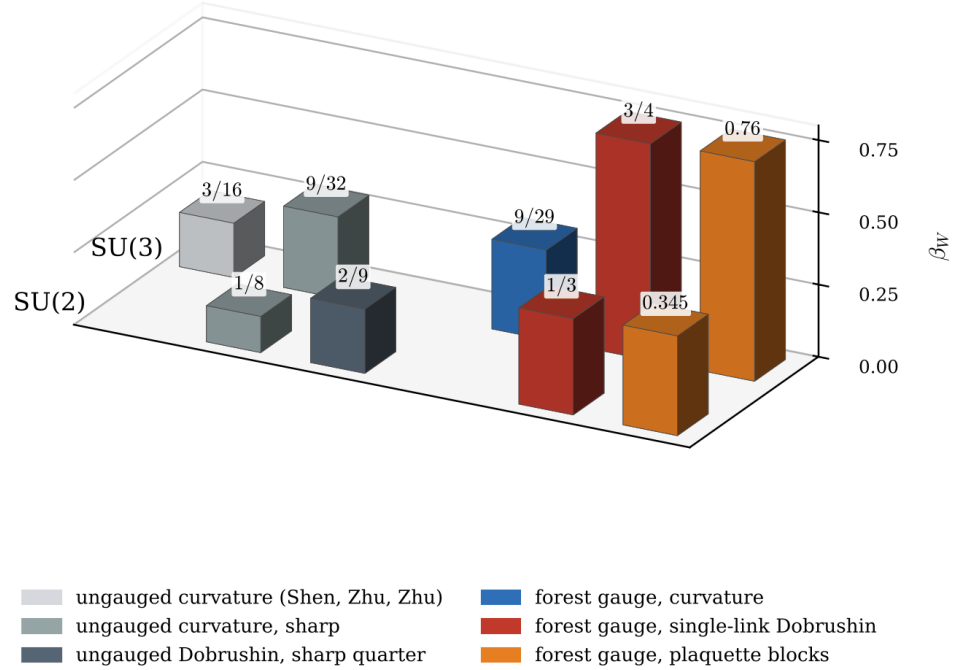


FIG. 2. Ranges of Table II. The height of each bar is the largest inverse coupling β_W covered by the method. Grey bars are methods for the ungauged measure; blue, red and orange bars use the forest gauge. For SU(2) the red bar ends at the ceiling of Theorem 2 and the orange bar, plaquette blocks, passes it (Theorem 3); for SU(3) the red bar is Theorem 6 and the orange bar, plaquette blocks, gives a larger gap there and passes it (Theorem 7).

pling, forces the gauge-fixed variables to be disordered on the scale of the cube, and cannot be reached from estimates on single lines of links; for SU(N) it needs cubes of side at least of order $(\beta_W / \log \beta_W)^{1/4}$ (Propositions 75 and 76). For a Gaussian field it holds at a fixed physical scale when the field has a mass and fails exactly, at every size, when it is massless (Proposition 62); it also fails for the quadratic approximation of the forest gauge itself (Proposition 63). The order-two cases of the tree bounds force the lattice gap below a fixed multiple of the lattice spacing, so no strong coupling expansion gives them (Proposition 59). The slice ratio holds in two dimensions, where it is the exact solution of lattice gauge theory (Remark 53), and fails for a field whose mass is fixed in lattice units; for the free field, conditioning an energy packet on averages over blocks leaves a fixed fraction of its ordinary variance at every lattice spacing, so a reduction to block averages has to be measured in the reflected norm (Proposition 54).

Block averages. Sec. XII poses the cube condition for block averages in a gauge whose fixed links are trees inside blocks of side two. Its quadratic approximation then holds at every coupling once the cube is large (Theorem 88), and integrating out one scale reduces it to a cube condition for the averages, which carry at least the coarse quadratic form at every scale (Corollary 91 and Proposition 89); the averages along the lines of the forest gauge, by contrast, leave one class of lines invisible and fail at every size (Proposition 82). For the group-valued links, the straight path between block roots gives a coarse link whose fibres satisfy a coercivity bound for every

compact group and every field (Theorem 94), but this map does not compose across scales (Proposition 97), and for every compact connected group no continuous covariant block map treats the positions of a block alike (Proposition 98). A cut-off of Balaban's averaging operation is continuous, covariant and coercive for every field and agrees with the block averages to first order; it breaks the symmetry only at large fields, as it must, and its constants are those of one step (Theorem 100). On fields with values in a maximal torus whose curvature stays small at every scale it composes, with the energy nonincreasing and the thin paths contracted (Theorem 102), and for every compact group it composes on fields whose curvatures at the successive scales have a small sum (Theorem 106). Under the SU(2) Wilson measure the plaquettes that are not small lie in clusters whose physical size tends to zero in probability, and the cut-off acts only next to them (Theorem 108, Corollaries 109 and 110); such plaquettes are nevertheless present along the one-loop tuning (Proposition 111). Whether some block map is coercive for all fields and composes across scales, through these clusters, is open.

Plan. Sec. II constructs the line field. Sec. III shows that a Dobrushin bound for the gauge-fixed specification gives the full conclusion of Theorem 1. Secs. IV, V, VI, VII and VIII prove Theorems 1–7, and Sec. IX measures what is left. Sec. X proves (I1)–(I3), Sec. XI states and measures their open inputs, and Sec. XII poses the cube condition for block averages, composes the cut-off average across scales at small fields and locates the large fields under the Wilson measure. The appen-

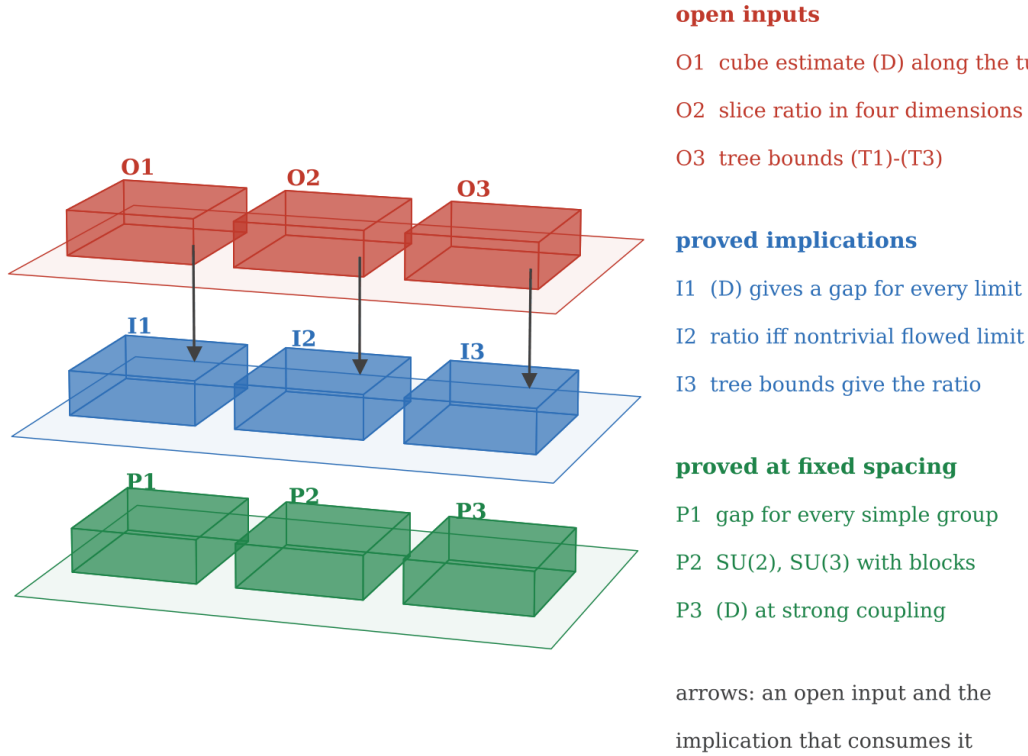


FIG. 3. The logical structure of the paper. Green: theorems at fixed lattice spacing (P1 the gap for every compact simple group, Theorem 5, P2 the plaquette blocks for SU(2) and SU(3), P3 the cube condition at strong coupling). Blue: the proved implications (I1)–(I3) of the continuum step. Red: their inputs, none of which is proved along the continuum tuning. Each arrow joins an open input to the implication that consumes it; no implication feeds its own input.

dices contain the residue table of the line field, the comparison lemma and its version for blocks, the blocks of the local Hessian inequality, the details of the certified computation, the reconstruction of the transfer matrix with the equivalence of a gap and exponential clustering (Appendix F), the facts on spectral measures and on the character expansion that the continuum step uses (Appendices G and H), and a complete proof of the sharp quarter for the single-link law of SU(2) (Appendices I–K), so that the paper does not depend on our earlier one for any step.

II. THE LINE FIELD

Write $e_0, \dots, e_3$ for the unit vectors of $\mathbb{Z}^4$ and (x, μ) for the link from x to $x + e_\mu$. A *direction field* is a map $\sigma : \mathbb{Z}^4 \rightarrow \{0, 1, 2, 3\}$; its links are $F = \{(x, \sigma(x)) : x \in \mathbb{Z}^4\}$. At each site put $\pi_x(\mu) = \sigma(x + e_\mu)$. The plaquette at x in the plane $\{\mu, \nu\}$ has the links (x, μ) , $(x + e_\mu, \nu)$, $(x + e_\nu, \mu)$, (x, ν) , so it contains

$$[\sigma(x) \in \{\mu, \nu\}] + [\pi_x(\mu) = \nu] + [\pi_x(\nu) = \mu] \quad (1)$$

links of F .

Lemma 8. *Every plaquette of $\mathbb{Z}^4$ contains exactly one link of F if and only if, at every site x , π_x fixes $\sigma(x)$ and permutes the other three directions cyclically.*

Proof. Fix x , let $s = \sigma(x)$ and let R be the set of the other three directions. For a plane $\{s, \nu\}$ the count (1) is $1 + [\pi_x(s) = \nu] + [\pi_x(\nu) = s]$; it equals one for all $\nu \in R$ exactly when $\pi_x(s) = s$ and $\pi_x(\nu) \neq s$ for $\nu \in R$, that is, when π_x fixes s and maps R into R . For a plane $\{\mu, \nu\} \subset R$ the count is $[\pi_x(\mu) = \nu] + [\pi_x(\nu) = \mu]$. Draw an arrow $\mu \rightarrow \pi_x(\mu)$ for every $\mu \in R$ with $\pi_x(\mu) \in R \setminus \{\mu\}$. Each vertex of R has at most one outgoing arrow, and the condition says that each of the three pairs of R carries exactly one arrow. Three arrows on three vertices with out-degree at most one give out-degree exactly one everywhere, and a tournament on three vertices with all out-degrees one is a directed three-cycle. $\square$

From now on σ is the field of period two given in Table III of Appendix A, read from x modulo two. The table lists π_x at each of the sixteen residue classes; in each row π_x fixes $\sigma(x)$ and is a three-cycle on the rest, so every plaquette contains exactly one link of F by Lemma 8.⁹

⁹ There are exactly eight period-two direction fields with this property, all

Lemma 9. (i) F is a disjoint union of bi-infinite straight lines, one through every site: every site has exactly one outgoing and one incoming link of F , and both point in the direction $\sigma(x)$.

(ii) For every box Λ (a product of four finite intervals, with the links and plaquettes it contains), $F \cap E(\Lambda)$ is a disjoint union of segments, hence a forest.

(iii) A link $e = (x, \mu) \notin F$ lies in six plaquettes. Four of them contain a link of F that shares an end point with e , and these lie in the planes $\{\mu, \sigma(x)\}$ and $\{\mu, \pi_x(\mu)\}$. In the other two, which lie in the remaining plane $\{\mu, \rho\}$, the link of F is opposite to e .

(iv) Every link outside F shares a plaquette with exactly twelve other links outside F , two in each of its plaquettes.

Proof. (i) Since π_x fixes $\sigma(x)$, $\sigma(x + e_{\sigma(x)}) = \sigma(x)$. Because σ has period two, $x - e_\mu \equiv x + e_\mu$, so also $\sigma(x - e_{\sigma(x)}) = \sigma(x)$, and the link $(x - e_{\sigma(x)}, \sigma(x))$ is in F . Conversely, if $(x - e_\nu, \nu) \in F$, then $\sigma(x - e_\nu) = \nu$, and straightness at $x - e_\nu$ gives $\sigma(x) = \nu$. So the incoming link is unique and continues the outgoing one.

(ii) A line meets a box in a segment, and distinct lines are disjoint.

(iii) The plaquettes of e are, for each $\nu \neq \mu$, the one at x and the one at $x - e_\nu$ in the plane $\{\mu, \nu\}$. A link of F sharing an end point with e is the outgoing or the incoming link of F at x or at $x + e_\mu$. At x both point in the direction $s = \sigma(x) \neq \mu$, and they lie in the plaquettes at x and at $x - e_s$ in the plane $\{\mu, s\}$. At $x + e_\mu$ both point in the direction $t = \pi_x(\mu)$, which lies in $R \setminus \{\mu\}$ because π_x is a three-cycle on R , and they lie in the plaquettes at x and at $x - e_t$ in the plane $\{\mu, t\}$. These four plaquettes are distinct because $s \neq t$. Each plaquette contains exactly one link of F , so in the two plaquettes of the third plane $\{\mu, \rho\}$ the link of F is a link of direction μ or ρ not touching e ; the links of direction ρ in such a plaquette touch e , so it is the opposite link.

(iv) Each of the six plaquettes contains e , one link of F and two other links, and two distinct links share at most one plaquette. $\square$

III. GAUGE FIXING AND THE FREE-BOUNDARY STATE

Let $G \subset U(N)$ be compact with Haar measure dU , and let Λ be a box. The Wilson measure with free boundary is

$$W_\Lambda(dQ) = \frac{1}{Z_\Lambda} \exp\left(N\beta \sum_{p \in P(\Lambda)} \text{Re Tr } Q_p\right) \prod_{e \in E(\Lambda)} dQ_e.$$

equivalent to σ under the symmetries of the lattice. A field with one fixed link at every site and one in every plaquette exists only in four dimensions: a link lies in $2(d-1)$ plaquettes and a site carries $d(d-1)/2$ of them, and $2(d-1) = d(d-1)/2$ forces $d = 4$. The statements of this section are also checked in the Lean certificate of the archive.

The measure is invariant under the gauge transformations $Q_{(x,y)} \mapsto g_x Q_{(x,y)} g_y^{-1}$ with $g \in G^\Lambda$. Call the links outside F active and write A for their set. For a function O of the links, O^F is O with every link of F set to the identity.

Lemma 10 (Gauge fixing). Let μ_Λ^F be the probability measure on $G^{A \cap E(\Lambda)}$ whose density with respect to Haar measure is proportional to $\exp(N\beta \sum_{p \in P(\Lambda)} \text{Re Tr } Q_p^F)$. For every gauge-invariant O on $G^{E(\Lambda)}$,

$$W_\Lambda(O) = \mu_\Lambda^F(O^F).$$

Proof. Each segment of $F \cap E(\Lambda)$ is a path $x_0, x_1, \dots, x_k$ along one direction. Given the variables on F , put $g_{x_0} = 1$ and $g_{x_{i+1}} = g_{x_i} Q_{(x_i, x_{i+1})}$ on every segment. The transformed field has $Q_e^g = 1$ on $F \cap E(\Lambda)$, and for fixed values on F the map $Q_A \mapsto Q_A^g$ is a product of left and right translations, which preserves Haar measure. Gauge invariance of the action and of O and the change of variables give the claim. $\square$

No leftover holonomy appears because $F \cap E(\Lambda)$ is a forest. The measures μ_Λ^F are finite-volume versions of the specification γ^F on G^A : for $e \in A$, γ_e^F resamples Q_e from the law with density proportional to $\exp(N\beta \sum_{p \ni e} \text{Re Tr } Q_p^F)$ given all other active links. No infinite-volume gauge transformation is performed; γ^F is a specification in its own right. Fix a metric δ on G of diameter D that is bi-invariant (for $SU(2)$ the chordal metric on S^3 , $D = 2$; for $SU(N)$ the Hilbert–Schmidt distance). The Kantorovich coefficient of e' on e is

$$C_{ee'} = \sup \frac{W_1(\gamma_e^F(\cdot | Q), \gamma_{e'}^F(\cdot | Q'))}{\delta(Q_e, Q_{e'})},$$

the supremum over pairs of configurations that differ only at e' . It vanishes unless e and e' share a plaquette.

Proposition 11. Suppose that every row and every column of $(C_{ee'})$ sums to at most $\alpha < 1$. Then:

(i) γ^F has exactly one Gibbs measure μ_F ;

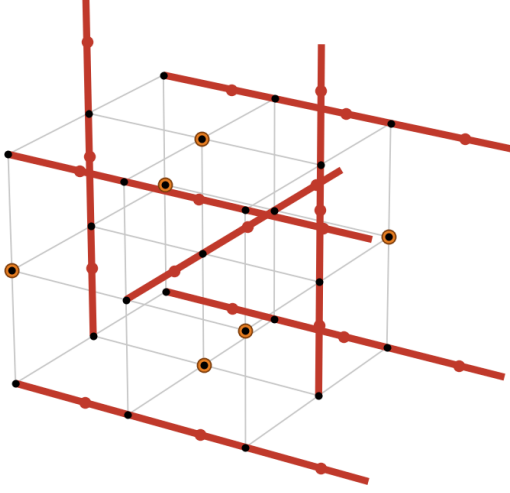
(ii) for every gauge-invariant local O the limit $\omega(O) = \lim W_\Lambda(O)$ exists along every sequence of boxes exhausting $\mathbb{Z}^4$ and equals $\mu_F(O^F)$;

(iii) ω is invariant under translations and the hypercubic group;

(iv) if $\beta \geq 0$, ω is reflection positive for reflections in hyperplanes through sites and in hyperplanes bisecting links;

(v) for gauge-invariant local O, O' with O Lipschitz in each link, $|\text{Cov}_\omega(O, O')| \leq C_{O, O'} e^{-\lambda \text{dist}(O, O')}$ for every $\lambda < -\log \alpha$;

(vi) if $\beta \geq 0$, the transfer matrix of ω has spectral gap at least $-\log \alpha$.

(a) the line field in the slice $x_3 = 0$ 

(b) one plaquette in the gauge

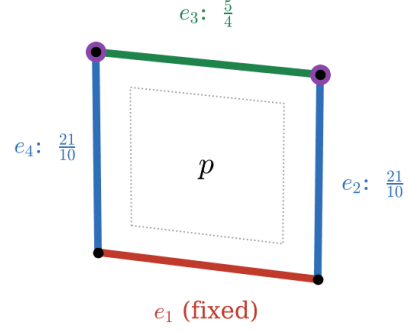


FIG. 4. (a) The line field in the slice $x_3 = 0$ of a box of side three. Thick red segments are the links of F in the directions 0, 1, 2, with a dot marking the direction of travel; an orange ring marks a site whose line leaves the slice in direction 3. Every line is straight (Lemma 9(i)), and the segments that leave the box are its boundary stubs. (b) One plaquette with its fixed link e_1 . The weights are those of the local inequality of Lemma 20 for the three free links: $\frac{21}{10}$ on the two links adjacent to e_1 and $\frac{5}{4}$ on the opposite one. The two purple corners are the ones that enter the corner form (7) once $Y_1 = 0$.

Proof. (i) Gibbs measures exist by compactness, and the comparison lemma of Appendix B with $S = \emptyset$ shows that any two coincide.

(ii) Extend μ_Λ^F by Haar measure on the active links outside Λ , and let S be the set of active links with a plaquette outside $P(\Lambda)$; this includes every active link outside $E(\Lambda)$. For an active link $e \notin S$ all six plaquettes of e lie in $P(\Lambda)$, so the conditional law of Q_e under the extended measure is γ_e^F . Let T be the support of O^F and $\lambda < -\log \alpha$. The comparison lemma with weights $e^{-\lambda \text{dist}(\cdot, T)}$ gives

$$|\mu_\Lambda^F(O^F) - \mu_F(O^F)| \leq L \frac{\theta D}{1 - \theta} \sum_{z \in S} e^{-\lambda \text{dist}(z, T)},$$

with $\theta = \alpha e^\lambda < 1$ and L a Lipschitz constant of O^F ; the sum is at most $C_T R^3 e^{-\lambda R}$ when R is the distance from T to the complement of Λ . For O continuous but not Lipschitz, approximate uniformly. Lemma 10 gives $W_\Lambda(O) = \mu_\Lambda^F(O^F)$.

(iii) $W_\Lambda(\tau_v O) = W_{\Lambda-v}(O)$, and the boxes $\Lambda - v$ also exhaust $\mathbb{Z}^4$; by (ii) the limit does not depend on the sequence. The line field is a device for computing W_Λ and is never shifted. The same argument applies to the hypercubic group.

(iv) For $\beta \geq 0$ the measure W_Λ is reflection positive for every box symmetric under the reflection [64], and reflection positivity passes to the limit along such boxes, which is ω by (ii).

(v) $\text{Cov}_\omega(O, O') = \text{Cov}_{\mu_F}(O^F, O'^F)$, and setting the links of F to the identity keeps the Lipschitz property. The bound is Corollary 113.

(vi) Let ϑ be a time reflection and τ_t a time translation. For gauge-invariant local O that are Lipschitz in each link, (v) gives $|\omega(\vartheta(\bar{O})\tau_t O) - \omega(O)|^2 \leq C_O e^{-\lambda t}$. On the support of O these

functions are dense in the gauge-invariant part of $L^2(\omega)$: gauge-invariant functions are functions on the compact quotient of G^E by the gauge group, Lipschitz functions for the quotient metric are dense in L^2 of every finite measure on it, and they lift to gauge-invariant Lipschitz functions. By Proposition 116 the spectrum of the transfer matrix on the complement of the vacuum lies in $[0, e^{-\lambda}]$, for every $\lambda < -\log \alpha$. $\square$

The proposition replaces the torus limits and the uniqueness of the ungauged Gibbs measure used in [14, Thm. F] by the free-boundary limit, which has all the properties the reconstruction of the transfer matrix needs.¹⁰

IV. SU(2) UP TO ONE THIRD

Proof of Theorem 1. Identify $\text{SU}(2)$ with the unit quaternions and use the chordal metric. For an active link e and a plaquette $p \ni e$, the staple s_p is the product of the other three link variables in the order that makes $\text{Re}(Q_e s_p) = \frac{1}{2} \text{Tr} Q_p$, with links of F equal to one. Since $N\beta \text{Re Tr} Q_p = 4\beta \langle Q_e, \bar{s}_p \rangle$ in $\mathbb{R}^4$, the conditional law of Q_e is the von Mises–Fisher law $\nu_b(dx) \propto e^{\langle x, b \rangle} dx$ on S^3 with $b = 4\beta \sum_{p \ni e} \bar{s}_p$. Changing a partner e' by δ in the chordal metric changes one staple by an isometric image of that displacement, so b moves by $4|\beta|\delta$,

¹⁰ On a torus the line field closes up into cycles, and the holonomies along them cannot be removed by a gauge transformation; the argument never uses a torus. The state ω is defined only on gauge-invariant observables, which is all that the Osterwalder–Schrader reconstruction [62–64] uses.

and the sharp quarter (Theorem 125), $W_1(v_b, v_{b'}) \leq \frac{1}{4}|b - b'|$, gives $C_{ee'} \leq |\beta|$. By Lemma 9(iv) every row and column has twelve nonzero entries, so $\alpha \leq 12|\beta| = 3|\beta_W|$. Proposition 11 applies for $|\beta_W| < \frac{1}{3}$. $\square$

Proof of Theorem 2. For the line field, every staple of an active link contains two active links, so the six staples can be chosen with sum zero while the staple through a given partner e' moves freely. At $b = 0$ the derivative of v_b in the direction h has Kantorovich norm $\frac{1}{4}|h|$ by Lemma 124(i), so $C_{ee'} = |\beta|$ exactly, and every row sum is $12|\beta| = 3|\beta_W|$.

Now let F' be any forest and C' the matrix of the statement. Suppose a positive weight w satisfies $C'w \leq \alpha w$. For a box Λ_n of side n let C_n be the principal submatrix on the active links of $E(\Lambda_n)$; then $C_n w \leq \alpha w$ on these links, and the Collatz–Wielandt inequality for nonnegative matrices gives $\rho(C_n) \leq \alpha$. Since C_n is symmetric, $\rho(C_n)$ is at least the mean row sum. The box has $4n^4 - O(n^3)$ links and n^4 sites, and a forest has fewer links than sites, so the number a of active links is at least $3n^4 - O(n^3)$. Let a_p be the number of active links of the plaquette p ; the box has $P = 6n^4 - O(n^3)$ plaquettes and $\sum_p a_p = 6a - O(n^3)$. Two links share at most one plaquette, so the row sums add up to $|\beta| \sum_p a_p (a_p - 1)$, and by convexity

$$\frac{|\beta|}{a} \sum_p a_p (a_p - 1) \geq \frac{|\beta|P}{a} \bar{a}(\bar{a} - 1) = 6|\beta|(\bar{a} - 1) - o(1),$$

with $\bar{a} = \sum_p a_p / P = 6a/P - o(1) \geq 3 - o(1)$. Hence $\alpha \geq 12|\beta| - o(1)$ for every n , and $\alpha \geq 12|\beta| = 3|\beta_W|$. $\square$

Without gauge fixing the same count gives eighteen partners and the ceiling $\beta_W = \frac{2}{9}$ of [14, Prop. 13]; the temporal gauge, a family of lines in one direction, leaves fourteen. The line field is optimal among forests: twelve is the mean number of partners forced by the number of free links, and it is attained at every link.

V. SU(2) PAST ONE THIRD: PLAQUETTE BLOCKS

Throughout this section $G = \text{SU}(2)$ is the group of unit quaternions with the chordal metric $|x - y|$, $\bar{x}$ is the conjugate, $\langle x, y \rangle = \text{Re}(x\bar{y})$ is the inner product of $\mathbb{R}^4$, and dx is the normalised Haar measure on S^3 . Left and right multiplication by unit quaternions and conjugation are isometries, and $\text{Re}(xyz) = \text{Re}(y\bar{z}x) = \langle y, \bar{z}\bar{x} \rangle$. Put

$$Z_0(c) = \int_{S^3} e^{\langle c, x \rangle} dx = \sum_{j \geq 0} \frac{(|c|^2/4)^j}{j!(j+1)!}, \quad c \in \mathbb{R}^4,$$

let $v_c(dx) \propto e^{\langle c, x \rangle} dx$ be the von Mises–Fisher law, and write $g = \log Z_0$, so that $\nabla g(c)$ is the mean and $\text{Hess } g(c)$ the covariance of v_c .

A. The blocks

Let $\mathcal{B}$ be the set of plaquettes at x in the plane $\{0, 1\}$ with $x_0 \equiv x_1$ and in the plane $\{2, 3\}$ with $x_2 \equiv x_3$, congruences

modulo two (Table IV of Appendix C lists them modulo $2\mathbb{Z}^4$, and Fig. 5 draws them). A *block plaquette* is a plaquette of $\mathcal{B}$, and its *block* is the set of its three active links.

Lemma 12. *The set $\mathcal{B}$ is invariant under translations by $2\mathbb{Z}^4$ and:*

- (i) *the blocks partition the active links;*
- (ii) *every plaquette $p \notin \mathcal{B}$ meets three distinct blocks, one in each of its active links;*
- (iii) *every active link lies in exactly five plaquettes outside $\mathcal{B}$, so it shares a plaquette outside $\mathcal{B}$ with exactly ten active links, none of them in its own block;*
- (iv) *if an active link j and a link of a block B lie in a common plaquette outside $\mathcal{B}$, then every link of B has its midpoint within distance 2 of the midpoint of j .*

Proof. Invariance under $2\mathbb{Z}^4$ is clear from the congruences. We use one fact throughout: two distinct links lie in at most one common plaquette, since they span it.

Every link lies in exactly one plaquette of $\mathcal{B}$. Indeed, a link (y, μ) with $\mu \in \{0, 1\}$ lies in no plaquette of the plane $\{2, 3\}$, and in exactly two of the plane $\{0, 1\}$, those at y and at $y - e_\nu$, where ν is the other element of $\{0, 1\}$. The base points y and $y - e_\nu$ differ only in the coordinate ν , so $y_0 - y_1$ has opposite parities at the two, and exactly one of them satisfies $x_0 \equiv x_1$. The directions 2, 3 are treated in the same way. Every plaquette of $\mathbb{Z}^4$ contains exactly one link of F (Lemma 8), so a block plaquette contains three active links, and every active link lies in exactly one block. This is (i).

(ii) Let $p \notin \mathcal{B}$ have active links e_1, e_2, e_3 , and let q_i be the block plaquette of e_i . Then $q_i \neq p$, so $q_i \cap p = \{e_i\}$ by the fact above. If $q_i = q_k$ with $i \neq k$, then e_i, e_k lie in the two distinct plaquettes p and q_i , which is impossible. So the three blocks are distinct, and each meets p in one link.

(iii) An active link e lies in six plaquettes (Lemma 9), and by (i) exactly one of them is in $\mathcal{B}$. Each of the other five contains two further active links. These ten links are distinct, because a link other than e shares at most one plaquette with e , and none lies in the block of e , because a link f of that block sharing the plaquette $p \notin \mathcal{B}$ with e would put e, f in the two distinct plaquettes p and the block plaquette of e .

(iv) Let $f \in B$ share the plaquette $p \notin \mathcal{B}$ with j . Two links of one unit square have midpoints at distance $2^{-1/2}$ or 1, so the midpoint of j is within distance 1 of that of f , and the midpoint of f is within distance 1 of that of every link of its block plaquette. The triangle inequality gives the bound 2. $\square$

Fix a block plaquette p with active links e_1, e_2, e_3 . Reading the plaquette product cyclically and deleting its fixed link, $\frac{1}{2} \text{Tr } Q_p = \text{Re}(y_1 y_2 y_3)$, where y_i is Q_{e_i} or $Q_{e_i}^{-1}$. Every other plaquette containing e_i meets the block only in e_i by Lemma 12(ii) and contributes $4\beta \langle y_i, \bar{s} \rangle$ to the exponent, where s is the staple of y_i in it. Hence the conditional law of the block given all other active links is

$$\mu_b(dy) \propto \exp\left(\sum_{i=1}^3 \langle y_i, b_i \rangle + 4\beta \text{Re}(y_1 y_2 y_3)\right) dy_1 dy_2 dy_3, \quad (2)$$

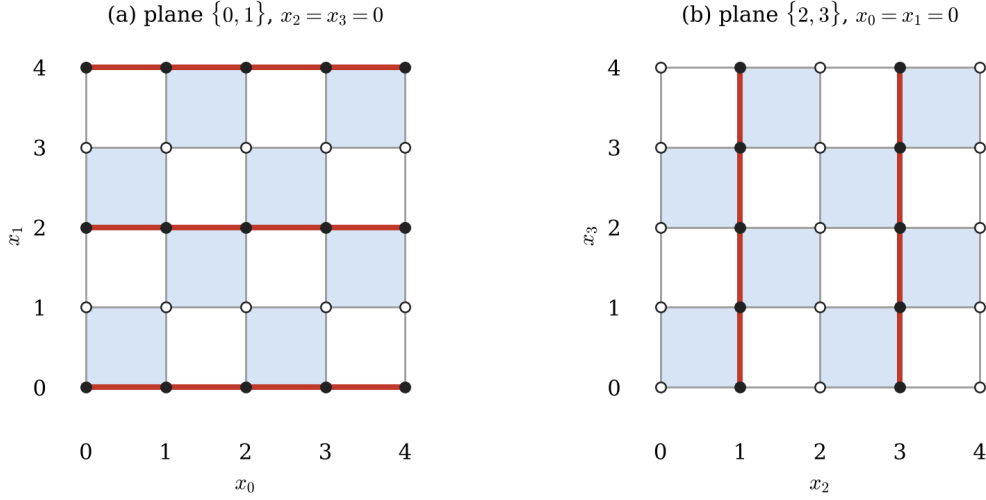


FIG. 5. Block plaquettes, shaded, in (a) a plane $\{0, 1\}$ and (b) a plane $\{2, 3\}$: the plaquettes at x with $x_0 \equiv x_1$, respectively $x_2 \equiv x_3$, modulo two. Red links belong to the line field F ; a filled dot marks a site whose line runs in the plane, an open dot one whose line leaves it. The shaded squares form a checkerboard, so every link of the plane borders exactly one of them, and every square contains exactly one red link (Lemma 8); hence each block plaquette carries three active links and the blocks partition the active links (Lemma 12(i)).

with b_i the sum of $4\beta\bar{s}$ over the five staples, so that $|b_i| \leq 20|\beta|$. Inversion is an isometry of S^3 and preserves Haar measure, so we may work with y instead of Q . On $(S^3)^3$ we use the metric $d(y, y') = \sum_i |y_i - y'_i|$.

B. The sensitivity of one block

Let λ_* and τ be a lower and an upper bound for the eigenvalues of $\text{Hess } g(c)$ on the ball $|c| \leq 24|\beta|$, and M an upper bound there for the norm of the third derivative $D^3g(c)$ as a trilinear form. Put

$$w = 8\beta^2(\tau - \lambda_*), \quad s = \frac{1}{4} + (e^w - 1)e^w\sqrt{\tau},$$

$$C = 20|\beta|(4|\beta|\tau s + 16\beta^2 M), \quad K = \frac{(1 + |\beta|)s}{1 - C}, \quad (3)$$

assuming $C < 1$.

Lemma 13. *Let ν be a probability measure and $\lambda = e^R \nu / \nu(e^R)$ with $\text{osc } R \leq \omega$. For bounded f, h ,*

$$\text{Cov}_\lambda(f, h) \leq \text{Cov}_\nu(f, h) + (e^\omega - 1)e^\omega \text{sd}_\nu(f) \text{sd}_\nu(h).$$

Proof. We may assume $0 \leq R \leq \omega$. The density $\rho = e^R / \nu(e^R)$ lies in $[e^{-\omega}, e^\omega]$, so $\delta = \|\rho - 1\|_\infty \leq e^\omega - 1$. With $\tilde{f} = f - \nu f$ and $\tilde{h} = h - \nu h$,

$$\begin{aligned} \text{Cov}_\lambda(f, h) - \text{Cov}_\nu(f, h) &= \nu((\rho - 1)\tilde{f}\tilde{h}) \\ &\quad - \nu((\rho - 1)\tilde{f}) \nu((\rho - 1)\tilde{h}), \end{aligned}$$

and the right side is at most $(\delta + \delta^2) \text{sd}_\nu(f) \text{sd}_\nu(h)$ by the Cauchy–Schwarz inequality. $\square$

Lemma 14. *Let $\beta \geq 0$, $a, c_0 \in \mathbb{R}^4$ with $|a|, |c_0| \leq 20\beta$, and*

$$\kappa(dx) \propto \exp(\langle x, a \rangle + g(c_0 + 4\beta\bar{x})) dx \quad \text{on } S^3.$$

For every 1-Lipschitz f and every $h \in \mathbb{R}^4$:

(i) $\text{Cov}_\kappa(f, \langle x, h \rangle) \leq s|h|;$

(ii) *if c_0 moves in the direction k , then $|\frac{d}{dt}\kappa(f)| \leq (4\beta\tau s + 16\beta^2 M)|k|$.*

Proof. (i) Taylor's formula at c_0 gives $g(c_0 + 4\beta u) = g(c_0) + 4\beta \langle \nabla g(c_0), u \rangle + R(u)$ with $R(u) = 8\beta^2 \text{Hess } g(\xi)(u, u)$ for some ξ on the segment, $|\xi| \leq 24\beta$. For $u = \bar{x}$ of norm one, $R(\bar{x}) \in [8\beta^2\lambda_*, 8\beta^2\tau]$, so $\text{osc } R \leq w$. Since $\langle \nabla g(c_0), \bar{x} \rangle = \langle x, \bar{\nabla} g(c_0) \rangle$, κ is the tilt $e^R \nu / \nu(e^R)$ of $\nu = \nu_c$ with $c = a + 4\beta \bar{\nabla} g(c_0)$. The sharp quarter, Theorem 125, gives $\text{Cov}_\nu(f, \langle x, h \rangle) = \frac{d}{dt} \nu_{c+th}(f) \leq \frac{1}{4}|h|$. A 1-Lipschitz function on a set of diameter two has $\text{sd} \leq 1$, and $\text{sd}_\nu \langle x, h \rangle \leq \sqrt{\tau}|h|$ because $|c| \leq |a| + 4\beta \leq 24\beta$. Lemma 13 gives (i).

(ii) The derivative is $\text{Cov}_\kappa(f, \langle \nabla g(c_0 + 4\beta\bar{x}), k \rangle)$. Write $\nabla g(c_0 + 4\beta u) = \nabla g(c_0) + 4\beta \text{Hess } g(c_0)u + N(u)$. The constant term does not contribute. The linear term is $\langle x, h \rangle$ with $h = 4\beta \text{Hess } g(c_0)k$, $|h| \leq 4\beta\tau|k|$, and contributes at most $4\beta\tau s|k|$ by (i). For $|u|, |u'| \leq 1$, $\|\text{Hess } g(c_0 + 4\beta u) - \text{Hess } g(c_0)\| \leq 4\beta M|u| \leq 4\beta M$ on the segment between them, so $|N(u) - N(u')| \leq 16\beta^2 M|u - u'|$; the function $\langle N(\bar{x}), k \rangle$ is therefore $16\beta^2 M|k|$ -Lipschitz, its standard deviation is at most that constant, and its covariance with f is at most $16\beta^2 M|k|$. $\square$

Lemma 15. *Let $C < 1$. If b and b' satisfy $|b_i|, |b'_i| \leq 20|\beta|$ and differ only in one coordinate b_i , then*

$$W_1(\mu_b, \mu_{b'}) \leq K|b_i - b'_i|$$

in the metric d .

Proof. The map $y_1 \mapsto -y_1$ turns μ_b for β into the law for $-\beta$ with b_1 replaced by $-b_1$ and is an isometry, so we may take

$\beta \geq 0$; the cyclic symmetry of $\text{Re}(y_1 y_2 y_3)$ lets us take $i = 1$. Let f be 1-Lipschitz for d and $\varepsilon = |b_1 - b'_1|$.

The middle link. Given y_1, y_3 , the law of y_2 is ν_c with $c = b_2 + 4\beta \bar{y}_3 \bar{y}_1$, which does not involve b_1 . Changing y_1 by η moves c by $4\beta \eta$, and the sharp quarter moves ν_c by at most $\beta \eta$ in W_1 . Hence $F(y_1, y_3) = \mathbb{E}[f | y_1, y_3]$ is $(1 + \beta)$ -Lipschitz in each variable, and $\mu_b(f) = \mu_b^{13}(F)$, where μ_b^{13} is the law of (y_1, y_3) :

$$\mu_b^{13} \propto e^{\langle y_1, b_1 \rangle + \langle y_3, b_3 \rangle} Z_0(b_2 + 4\beta \bar{y}_3 \bar{y}_1).$$

Since $|b_2 + 4\beta \bar{y}_3 \bar{y}_1| = |b_2 y_3 + 4\beta \bar{y}_1| = |y_1 b_2 + 4\beta \bar{y}_3|$, its conditional laws are

$$\begin{aligned} \kappa_1(dy_1 | y_3) &\propto e^{\langle y_1, b_1 \rangle + g(b_2 y_3 + 4\beta \bar{y}_1)} dy_1, \\ \kappa_3(dy_3 | y_1) &\propto e^{\langle y_3, b_3 \rangle + g(y_1 b_2 + 4\beta \bar{y}_3)} dy_3, \end{aligned}$$

both of the form of Lemma 14, with $|c_0| = |b_2|$.

Two sites. Integrating Lemma 14(i) along the segment from b_1 to b'_1 gives $W_1(\kappa_1, \kappa'_1) \leq s\varepsilon$ for the same y_3 . Moving y_3 to y'_3 moves $c_0 = b_2 y_3$ along a segment of length $|b_2||y_3 - y'_3| \leq 20\beta|y_3 - y'_3|$ inside the ball $|c_0| \leq 20\beta$, so Lemma 14(ii) gives $W_1(\kappa_1(\cdot | y_3), \kappa_1(\cdot | y'_3)) \leq C|y_3 - y'_3|$, and the same for κ_3 , which does not depend on b_1 . Let π be a coupling of μ_b^{13} and $\mu_{b'}^{13}$ minimising $d_1 + d_3$, $d_k = \mathbb{E}_\pi |y_k - y'_k|$. Replacing (y_1, y'_1) by an optimal coupling of $\kappa_1(\cdot | y_3)$ and $\kappa'_1(\cdot | y'_3)$ keeps a coupling, so minimality gives $d_1 \leq s\varepsilon + Cd_3$; in the same way $d_3 \leq Cd_1$. Hence $d_1 \leq s\varepsilon/(1 - C^2)$ and

$$|\mu_b(f) - \mu_{b'}(f)| \leq (1 + \beta)(d_1 + d_3) \leq \frac{(1 + \beta)s}{1 - C} \varepsilon = K\varepsilon. \quad \square$$

C. The constants

Lemma 16. Write $g(c) = \varphi(|c|^2)$ and $r = |c|$. For $c \neq 0$ the eigenvalues of $\text{Hess } g(c)$ are $2\varphi'(r^2)$, three times, and $2\varphi'(r^2) + 4r^2\varphi''(r^2)$, and the norm of $D^3 g(c)$ is

$$\max\left(|A + \Gamma|, \frac{2}{3}|A|\right), \quad A = 12r\varphi''(r^2), \quad \Gamma = 8r^3\varphi'''(r^2).$$

On the ball $r \leq 2.07$ every eigenvalue of $\text{Hess } g$ lies between 0.158491 and 0.25000001, and the norm of $D^3 g$ is at most 0.0615586.

Proof. Differentiating $\varphi(|c|^2)$ gives $D^2 g(v, v) = 2\varphi'|v|^2 + 4\varphi''\langle c, v \rangle^2$ and $D^3 g(v, v, v) = 12\varphi''\langle c, v \rangle|v|^2 + 8\varphi''' \langle c, v \rangle^3$. For $|v| = 1$ and $t = \langle c/r, v \rangle \in [-1, 1]$ the cubic form is $At + \Gamma t^3$, whose largest absolute value is attained at $t = \pm 1$ or at an interior critical point, where it equals $\frac{2}{3}|A||t| \leq \frac{2}{3}|A|$; by Banach's theorem [8] the supremum of a symmetric trilinear form on the diagonal of the unit sphere is its norm. The numerical bounds are obtained in ball arithmetic [48] from the series of Z_0 , truncated after forty terms with a proved bound on the tail, on $2 \cdot 10^4$ subintervals of $[0, 2.07]$; the program is described in Appendix C. $\square$

D. The proof of Theorem 3

Let $\gamma^\mathcal{B}$ be the block specification: for a block B , $\gamma_B^\mathcal{B}$ re-samples the three links of B from (2). It comes from the same interaction as γ^F , so the two specifications have the same Gibbs measures. Define the block coefficients

$$C_{Bj} = \sup \frac{W_1(\gamma_B^\mathcal{B}(\cdot | \mathcal{Q}), \gamma_B^\mathcal{B}(\cdot | \mathcal{Q}'))}{|\mathcal{Q}_j - \mathcal{Q}'_j|}$$

for an active link $j \notin B$, the supremum over pairs of configurations that differ only at j , with W_1 for the metric d .

Proposition 17. Suppose that $\sum_B C_{Bj} \leq \alpha < 1$ for every active link j . Then the conclusions (i)–(vi) of Proposition 11 hold with $-\frac{1}{2} \log \alpha$ in place of $-\log \alpha$.

Proof. The proof of Proposition 11 goes through with Lemma 114 of Appendix B in place of Lemma 112, and with the set S of part (ii), and the support of H in Corollary 113, enlarged to the union of the blocks that meet it. The enlargement multiplies the size of these sets by at most three and moves them by at most one, which changes only the constants. $\square$

Proof of Theorem 3. Let $|\beta| \leq 0.345/4$ and take the constants of Lemma 16, whose ball $r \leq 2.07$ contains $|c| \leq 24|\beta|$. With these fixed constants every quantity in (3) is increasing in $|\beta|$, and so is

$$\alpha(\beta_W) = 40|\beta|K. \quad (4)$$

Changing an active link j by η changes one staple of each of its ten partners outside its block by an isometric image of the displacement, so the field b_i of each such partner i moves by at most $4|\beta|\eta$, and Lemma 15, applied once for each of these ten fields, gives $\sum_B C_{Bj} \leq 40|\beta|K$. Evaluating (3) in ball arithmetic at $\beta_W = 0.345$ gives $w \leq 0.0054459$, $s \leq 0.252746$, $C \leq 0.050243$, $K \leq 0.289069$ and $\alpha \leq 0.99729$; at $\beta_W = \frac{1}{3}$ it gives $\alpha \leq 0.95602$. Proposition 17 applies. $\square$

Remark 18. Most of the range is kept. A second bound for the same sensitivity, which keeps the law of the middle link instead of linearising it, evaluated in floating point on a quadrature grid and maximised over the fields b by a random search followed by the Nelder–Mead method, gives 0.2783 at $\beta_W = 0.34$, against 0.2881 from Lemma 15, and $40|\beta|$ times it crosses one near $\beta_W = 0.357$. That computation is not certified and is not used in the proof. Past this point the gain from integrating out one link is spent, and larger blocks are needed, such as the free links of a cube of the line field, whose sensitivity no longer reduces to two sites.

VI. CURVATURE IN THE GAUGE

In this section $G \subset \text{U}(N)$ is a closed connected subgroup with Lie algebra $\mathfrak{g}$ and the inner product $\langle X, Y \rangle = -\text{Tr}(XY)$; the bound of Theorem 4 is proved first for $\text{U}(N)$. A tangent vector at a field Q is $v = (X_e Q_e)$, $|v|^2 = \sum_e |X_e|^2$. For the plaquette at x in the plane $\mu < \nu$ number its links $e_1 = (x, \mu)$,

$e_2 = (x + e_\mu, v)$, $e_3 = (x + e_v, \mu)$, $e_4 = (x, v)$, transport their vectors to x ,

$$Y_1 = X_{e_1}, \quad Y_2 = Q_{e_1} X_{e_2} Q_{e_1}^*, \quad Y_3 = Q_{e_4} X_{e_3} Q_{e_4}^*, \quad Y_4 = X_{e_4}, \quad (5)$$

and write $W = Q_p$. The second derivative of $\text{Re Tr } Q_p$ along the geodesic $Q_e(t) = e^{tX_e} Q_e$ is

$$\begin{aligned} f_p &= \sum_{i=1}^4 \text{Re Tr}(Y_i^2 W) + 2 \text{Re Tr}(Y_1 Y_2 W) + 2 \text{Re Tr}(Y_3 Y_4 W) \\ &\quad - 2 \sum_{i \in \{1,2\}, j \in \{3,4\}} \text{Re Tr}(Y_i W Y_j), \end{aligned} \quad (6)$$

so that $\text{Hess } S_0(v, v) = \sum_p f_p$. To see this, write $Q_p = Q_{e_1} Q_{e_2} Q_{e_3}^* Q_{e_4}^*$. Since $Q_{e_1} e^{tX_{e_2}} = e^{tY_2} Q_{e_1}$ and $e^{-tX_{e_3}} Q_{e_4}^* = Q_{e_4}^* e^{-tY_3}$, the plaquette variable along the geodesic is $Q_p(t) = e^{tY_1} e^{tY_2} W e^{-tY_3} e^{-tY_4}$. Differentiating twice at $t = 0$ gives the four squares, the two products of neighbouring factors and the four products across W , and the squares and the product $W Y_3 Y_4$ are moved cyclically under the trace. The corner form of p is

$$c_p = 2\langle Y_1, Y_4 \rangle - 2\langle Y_1, Y_2 \rangle + 2\langle Y_2, W Y_3 W^* \rangle - 2\langle Y_3, Y_4 \rangle. \quad (7)$$

It collects the inner products of the transported vectors of two links that meet at a right angle at a corner of p .

Lemma 19. *On $\mathbb{Z}^4$, or on a box with free boundary, for every field Q and every square-summable X ,*

$$\begin{aligned} &\sum_y \left[|D_y X|^2 + \sum_\mu |X_{y,\mu} + Q_{y-e_\mu,\mu}^* X_{y-e_\mu,\mu} Q_{y-e_\mu,\mu}|^2 \right] \\ &= 4|X|^2 + \sum_p c_p, \end{aligned}$$

where $D_y X = \sum_\mu [X_{y,\mu} - Q_{y-e_\mu,\mu}^* X_{y-e_\mu,\mu} Q_{y-e_\mu,\mu}]$ is the covariant divergence and terms with a link outside the box are omitted.

Proof. At a site y write o_μ for the vector of the outgoing link in direction μ and i_μ for the transported vector of the incoming one, so that $D_y X = \sum_\mu (o_\mu - i_\mu)$. Every link is outgoing at one end and incoming at the other, so the squares $|o_\mu|^2 + |i_\mu|^2$, which appear once in $|D_y X|^2$ and once in $|o_\mu + i_\mu|^2$, add up to $4|X|^2$. The collinear cross term $-2\langle o_\mu, i_\mu \rangle$ of $|D_y X|^2$ cancels the term $+2\langle o_\mu, i_\mu \rangle$ of $|o_\mu + i_\mu|^2$. The remaining cross terms of $|D_y X|^2$ pair two perpendicular ends at y , and each such pair is a corner of exactly one plaquette, with the sign and the transport of (7) (this is the proof of [14, Lemma 1]). In a box two perpendicular ends at y span a plaquette of the box, because a box is a product of intervals. $\square$

Lemma 20. *Let $k \in \{1, 2, 3, 4\}$, let $k \pm 1 \pmod{4}$ be the positions adjacent to k and $k + 2$ the opposite one. For every N , every $W \in \text{U}(N)$ and all $Y_1, \dots, Y_4 \in \text{u}(N)$ with $Y_k = 0$,*

$$\pm f_p + \frac{9}{10} c_p \leq \frac{21}{10} (|Y_{k-1}|^2 + |Y_{k+1}|^2) + \frac{5}{4} |Y_{k+2}|^2.$$

Proof. Replacing W by $-W$ changes the sign of f_p and fixes c_p , so the sign $+$ suffices. Conjugation by a unitary matrix preserves both sides, so we may take $W = \text{diag}(w_a)$. The form $\mathcal{F} = \text{RHS} - f_p - \frac{9}{10} c_p$ then splits into a block for each diagonal position a , in the real variables $y_i = -i(Y_i)_{aa}$, and a block for each pair $a \neq b$, in the complex variables $z_i = (Y_i)_{ab}$, since $(Y_i)_{ba} = -\bar{z}_i$. Appendix D shows that every block is positive definite. $\square$

Proof of the bound in Theorem 4. Every plaquette has exactly one link of F , and X vanishes on it. By Lemma 9(iii) each active link is adjacent to the fixed link in four of its plaquettes and opposite to it in two. Summing Lemma 20 over the plaquettes,

$$\sum_p (\pm f_p + \frac{9}{10} c_p) \leq (4 \cdot \frac{21}{10} + 2 \cdot \frac{5}{4}) |X|^2 = \frac{109}{10} |X|^2.$$

By Lemma 19, $\frac{9}{10} \sum_p c_p = \frac{9}{10} \Sigma - \frac{36}{10} |X|^2$, where $\Sigma \geq 0$ is the left side of the lemma. Hence

$$\pm \text{Hess } S_0(v, v) + \frac{9}{10} \Sigma \leq (\frac{109}{10} + \frac{36}{10}) |X|^2 = \frac{29}{2} |X|^2.$$

In a box, links near the boundary lie in fewer plaquettes and receive less. $\square$

The factor $\frac{9}{10}$ is chosen so that the corners are paid by the divergence identity at the price $\frac{36}{10}$ while the local blocks stay positive; Lemma 19 is the divergence identity of [14, Lemma 1] with the collinear term written as a square, so nothing is lost by Cauchy–Schwarz.

The two lemmas concern vectors in $\text{u}(N)$ and fields in $\text{U}(N)$, so the bound holds without change on every closed subgroup $G \subset \text{U}(N)$, the vectors X_e ranging over its Lie algebra $\mathfrak{g}$. It is also the Hessian of S_0 on G^E : the metric is bi-invariant, the geodesics of G through Q_e are the curves $e^{tX} Q_e$ with $X \in \mathfrak{g}$, which are geodesics of $\text{U}(N)$, and the Hessian is read off from second derivatives along geodesics.

Lemma 21. *Let $G \subset \text{U}(N)$ be closed and connected with simple Lie algebra $\mathfrak{g}$, let $\mathfrak{t} \subset \mathfrak{g}$ be a maximal abelian subalgebra, and write the eigenvalues of $\text{ad } H$ on $\mathfrak{g} \otimes \mathbb{C}$ and of H on $\mathbb{C}^N$, for $H \in \mathfrak{t}$, as $i\alpha(H)$ and $i\mu(H)$, with α running over the roots and μ over the weights of $\mathbb{C}^N$ counted with multiplicity. Then $\text{Ric}_G = \kappa_G \langle \cdot, \cdot \rangle$ with*

$$\kappa_G = \frac{\sum_\alpha \alpha(H)^2}{4 \sum_\mu \mu(H)^2} \quad (0 \neq H \in \mathfrak{t}).$$

Proof. For a bi-invariant metric the curvature is $R(X, Y)Z = -\frac{1}{4} [[X, Y], Z]$ on left-invariant fields [60, Sect. 7], so for an orthonormal basis (E_i) of $\mathfrak{g}$,

$$\begin{aligned} \text{Ric}(X, X) &= \sum_i \langle R(E_i, X) X, E_i \rangle \\ &= \frac{1}{4} \sum_i |[X, E_i]|^2 = -\frac{1}{4} \text{Tr}(\text{ad } X)^2, \end{aligned}$$

since $\text{ad } X$ is skew. The Killing form $B(X, Y) = \text{Tr}(\text{ad } X \text{ ad } Y)$ and $\langle \cdot, \cdot \rangle$ are invariant under the adjoint action, and $\mathfrak{g}$ is simple, so $B = -4\kappa \langle \cdot, \cdot \rangle$ for a constant κ and $\text{Ric} = \kappa \langle \cdot, \cdot \rangle$. On $H \in \mathfrak{t}$ the two forms are $B(H, H) = -\sum_\alpha \alpha(H)^2$ and $\langle H, H \rangle = -\text{Tr } H^2 = \sum_\mu \mu(H)^2$. $\square$

Table I evaluates the ratio of Lemma 21 for every compact simple Lie algebra and a faithful representation of least dimension, on the root systems of [46, Sect. 12.1]; the weights of the representations of dimensions 27 and 56 of E_6 and E_7 are the roots of E_8 with prescribed inner products with a subsystem A_2 , respectively A_1 , projected to its orthogonal complement, by the decompositions $248 = (78, 1) + (27, 3) + (\bar{27}, \bar{3}) + (1, 8)$ and $248 = (133, 1) + (56, 2) + (1, 3)$ [70]. Taking for H the coroot $\theta^\vee$ of the highest root gives $\sum_\alpha \alpha(\theta^\vee)^2 = 4h^\vee$ and $\sum_\mu \mu(\theta^\vee)^2 = 2\ell_\rho$, so $\kappa_G = h^\vee/(2\ell_\rho)$, with the Dynkin index normalised to one on the defining representation of $SU(N)$. The ratio depends on $\mathfrak{g}$ and ρ only, so it is the same for every compact group with Lie algebra $\mathfrak{g}$, simply connected or not.

Lemma 22. *Let $e \neq e'$ be two links of a plaquette p , and write*

$$\text{Hess Re Tr } Q_p(v, v) = \sum_{e, e' \in p} H_{ee'}^p(X_e, X_{e'})$$

with bilinear forms satisfying $H_{ee'}^p(a, b) = H_{e'e}^p(b, a)$. Then $|H_{ee'}^p(a, b)| \leq |a| |b|$ for all $a, b \in \mathfrak{u}(N)$ and every field.

Proof. By (6) the terms of f_p that mix two positions $i \neq j$ are $2\text{Re Tr}(Y_1 Y_2 W)$, $2\text{Re Tr}(Y_3 Y_4 W)$ and $-2\text{Re Tr}(Y_i W Y_j)$, so $H_{ee'}^p$ is $\pm \text{Re Tr}(abW)$ or $\pm \text{Re Tr}(aWb)$ after an isometric transport of a and b , and $|\text{Tr}(abW)| \leq \|a\|_{\text{HS}} \|bW\|_{\text{HS}} = |a| |b|$. $\square$

Corollary 23. *Let $G \subset U(N)$ be closed and connected with semisimple Lie algebra and $N|\beta| < \frac{2}{29}\kappa_G$. For every box the measure μ_Λ^F satisfies the Bakry–Émery condition with $K = \kappa_G - \frac{29}{2}N|\beta| > 0$, uniformly in the box, and hence a logarithmic Sobolev inequality with constant $2/K$ and a Poincaré inequality with constant $1/K$.*

Proof. The set $\{Q_F = 1\}$ is a product factor, so its Ricci curvature is at least κ_G on every link, and the Hessian of the restricted action is the restriction of $N\beta \text{Hess } S_0$ to $X_F = 0$. Apply the bound and [4, Cor. 5.7.2]. $\square$

A. Decay with an explicit rate

Let A' be a finite set of active links and P' a finite set of plaquettes whose active links lie in A' . An action $S = \sum_{p \in P'} c_p \text{Re Tr } Q_p^F$ with $0 \leq c_p \leq 1$ is admissible if $|\text{Hess } S(v, v)| \leq \frac{29}{2}|v|^2$ for all v tangent to $G^{A'}$. Two links lie in at most one common plaquette, so by Lemma 22 and Lemma 9(iv) the mixed blocks of its Hessian satisfy $|\text{Hess } S_{ee'}(a, b)| \leq |a| |b|$, and each link has at most twelve partners. The action $S_\Lambda = \sum_{p \in P(\Lambda)} \text{Re Tr } Q_p^F$ of a box is admissible by the bound proved above, and so is every convex combination of two admissible actions. Let μ be the probability measure on $G^{A'}$ with density proportional to $e^{N\beta S}$ with respect to Haar measure, let $L = \Delta + N\beta \langle \nabla S, \nabla \cdot \rangle$ be its Langevin generator, which is self-adjoint in $L^2(\mu)$, and $P_t = e^{tL}$. Write ∇_e for the gradient in the factor of the link e , and for weights $w : A' \rightarrow (0, \infty)$ put $\Gamma^w(f) = \sum_e w_e |\nabla_e f|^2$; Γ is the case $w \equiv 1$.

Lemma 24. *Let $\lambda \geq 0$ and $w_e = e^{\lambda d_e}$, where $|d_e - d_{e'}| \leq 1$ whenever e and e' share a plaquette, and put*

$$K_\lambda = \kappa_G - N|\beta| \left(\frac{29}{2} + 12(\cosh \frac{\lambda}{2} - 1) \right).$$

Then $\Gamma^w(P_t f) \leq e^{-2K_\lambda t} P_t \Gamma^w(f)$ for every smooth f and $t \geq 0$.

Proof. Put $V = -N\beta S$. On the product manifold $M = G^{A'}$ the projection to the tangent space of one factor is parallel, the rough Laplacian commutes with it, and the Ricci tensor is block diagonal. Hence the Bochner–Weitzenböck formula $\nabla \Delta g = \bar{\Delta} \nabla g + \text{Ric}(\nabla g)$, applied factor by factor, together with the derivative of the drift, gives for smooth g

$$\begin{aligned} \frac{1}{2} L \Gamma^w(g) - \Gamma^w(g, Lg) &= \sum_e w_e (|\nabla \nabla_e g|^2 + \text{Ric}_G(\nabla_e g, \nabla_e g)) \\ &\quad + \sum_{e, e'} w_e \text{Hess } V_{ee'}(\nabla_e g, \nabla_{e'} g), \end{aligned}$$

where $\Gamma^w(g, h)$ is the polarised form.¹¹ Put $u_e = \nabla_e g$ and $v_e = \sqrt{w_e} u_e$. By the symmetry of $\text{Hess } V$,

$$\begin{aligned} \sum_{e, e'} w_e \text{Hess } V_{ee'}(u_e, u_{e'}) &= \text{Hess } V(v, v) \\ &\quad + \sum_{e \neq e'} \left(\cosh \frac{\lambda(d_e - d_{e'})}{2} - 1 \right) \text{Hess } V_{ee'}(v_e, v_{e'}), \end{aligned}$$

because $\frac{w_e + w_{e'}}{2\sqrt{w_e w_{e'}}} = \cosh \frac{\lambda(d_e - d_{e'})}{2}$. The first term is at least $-\frac{29}{2}N|\beta| |v|^2$. In the second only partners contribute, each with a coefficient at most $\cosh \frac{\lambda}{2} - 1$ and a block of norm at most $N|\beta|$, and $2|v_e| |v_{e'}| \leq |v_e|^2 + |v_{e'}|^2$ with twelve partners per link bounds it by $12(\cosh \frac{\lambda}{2} - 1)N|\beta| |v|^2$. Since $\text{Ric}_G \geq \kappa_G$ and $|v|^2 = \Gamma^w(g)$,

$$\frac{1}{2} L \Gamma^w(g) - \Gamma^w(g, Lg) \geq K_\lambda \Gamma^w(g).$$

For fixed t let $\varphi(s) = P_s \Gamma^w(P_{t-s} f)$, $0 \leq s \leq t$. Then $\varphi'(s) = P_s (L \Gamma^w(g_s) - 2\Gamma^w(g_s, Lg_s)) \geq 2K_\lambda \varphi(s)$ with $g_s = P_{t-s} f$, and $\varphi(0) \leq e^{-2K_\lambda t} \varphi(t)$ is the claim. $\square$

Lemma 25. *Suppose $K_\lambda > 0$. Let f and g be smooth functions on $G^{A'}$ that depend only on the links of A_f and of B_g , and let D be the least number of steps from A_f to B_g in the graph on A' whose edges join links that share a plaquette. Then*

$$|\text{Cov}_\mu(f, g)| \leq \frac{e^{-\lambda D/2}}{K_\lambda} \|\nabla f\|_\infty \sum_{e \in B_g} \|\nabla_e g\|_\infty,$$

where $\|\nabla f\|_\infty^2 = \sup \sum_e |\nabla_e f|^2$.

¹¹ For $w \equiv 1$ this is Bochner's identity $\Gamma_2 = |\nabla^2 g|^2 + (\text{Ric} + \text{Hess } V)(\nabla g, \nabla g)$; the weights enter only through the factor index, which the connection, the Laplacian and the Ricci tensor preserve, so only the off-diagonal blocks of $\text{Hess } V$ see them.

Proof. Since $K_0 \geq K_\lambda > 0$, Lemma 24 with $\lambda = 0$ gives $\Gamma(P_t f) \leq e^{-2K_0 t} \|\nabla f\|_\infty^2$, so $P_t f$ converges uniformly to a constant, which is $\mu(f)$ by invariance, M being connected. Since $\frac{d}{dt} \mu(g P_t f) = \mu(g L P_t f) = -\mu(\Gamma(g, P_t f))$,

$$\text{Cov}_\mu(f, g) = \int_0^\infty \mu(\Gamma(P_t f, g)) dt.$$

Let d_e be the smaller of D and the graph distance from e to A_f , so that $|d_e - d_{e'}| \leq 1$ for partners, $w_e = 1$ on A_f , $w_e = e^{\lambda D}$ on B_g , and $\Gamma^w(f) \leq \|\nabla f\|_\infty^2$. For $e \in B_g$, Lemma 24 gives $|\nabla_e P_t f|^2 \leq w_e^{-1} \Gamma^w(P_t f) \leq e^{-\lambda D} e^{-2K_\lambda t} \|\nabla f\|_\infty^2$. Hence $|\Gamma(P_t f, g)| \leq e^{-\lambda D/2} e^{-K_\lambda t} \|\nabla f\|_\infty \sum_{e \in B_g} \|\nabla_e g\|_\infty$, and integration in t gives the bound. $\square$

Proof of Theorem 5. Let $\lambda^* > 0$ solve $K_{\lambda^*} = 0$, so that $\frac{\lambda^*}{2} = m_G(\beta)$, and let $0 < \lambda < \lambda^*$; then $K_\lambda > 0$.

Decay in boxes. Lemma 25 with $S = S_\Lambda$ gives

$$|\text{Cov}_{\mu_\Lambda^F}(f, g)| \leq C_{f,g} e^{-\lambda D/2},$$

uniformly in the box.

Limit. For boxes $\Lambda \subset \Lambda'$ let $S_s = (1-s)S_\Lambda + sS_{\Lambda'}$ on the active links of Λ' , $s \in [0, 1]$, and μ_s the corresponding measure; S_s is admissible, μ_0 is μ_Λ^F times Haar measure on the active links outside Λ , and $\mu_1 = \mu_{\Lambda'}^F$. For smooth f of the active links of Λ ,

$$\frac{d}{ds} \mu_s(f) = N\beta \sum_{p \in P(\Lambda') \setminus P(\Lambda)} \text{Cov}_{\mu_s}(f, \text{Re Tr } Q_p^F),$$

and $|\nabla_e \text{Re Tr } Q_p| \leq \|Q_p\|_{\text{HS}} = \sqrt{N}$. A step in the graph of Lemma 25 moves a link by at most one lattice unit, so a plaquette at graph distance k from the support T of f lies within distance $k+1$ of T , there are at most $C_T(k+1)^4$ of them, and every plaquette outside $P(\Lambda)$ is at graph distance at least $R-1$, R the distance from T to the complement of Λ . Lemma 25 gives

$$|\mu_{\Lambda'}^F(f) - \mu_\Lambda^F(f)| \leq C_f R^4 e^{-\lambda R/2}.$$

The limit therefore exists along every increasing sequence of boxes, and two sequences are compared through their inter-sections, which are boxes. Continuous gauge-invariant O are uniform limits of smooth gauge-invariant ones, for instance of $e^{\tau \Delta} O$, since the heat semigroup of G^E commutes with left and right translations; with Lemma 10 the limit $\omega(O)$ of $W_\Lambda(O)$ exists for all of them.

Properties. Invariance and, for $\beta \geq 0$, reflection positivity follow as in Proposition 11(iii)–(iv). For Lipschitz O the functions $e^{\tau \Delta} O$ are smooth with the same Lipschitz constant, since $\text{Ric}_G \geq 0$ [4, Sect. 3.3], so the bound of the first step passes to ω and to Lipschitz observables. Two links of one plaquette have time coordinates that differ by at most one, so observables at time separation t are at graph distance at least $t-c$, and truncated correlations decay at every rate below $m_G(\beta)$. The gap follows from Proposition 116 as in Proposition 11(vi). $\square$

Proof of the rest of Theorem 4. By Lemma 21 and Table I, $\kappa_{\text{SU}(N)} = \frac{N}{2}$ and $\kappa_{\text{SO}(N)} = \frac{N-2}{4}$; for $\text{SO}(4)$, whose Lie algebra is $\mathfrak{so}(3) \oplus \mathfrak{so}(3)$ acting on $\mathbb{R}^4 \otimes \mathbb{C} = \mathbb{C}^2 \otimes \mathbb{C}^2$, each ideal gives $\frac{1}{2}$ by the same computation. The conditions of Theorem 4 are those of Theorem 5. $\square$

Remark 26. Let h_F be the supremum of $\pm \text{Hess } S_0$ on the gauge-fixed directions, so that uniform curvature reaches $|\beta| < 1/(2h_F)$ and no further. At the identity field $-\text{Hess } S_0$ restricted to $X_F = 0$ has the eigenvalue $9 + \sqrt{17} = 13.123$ on the torus of side two, so $9 + \sqrt{17} \leq h_F \leq \frac{29}{2}$. A gradient search over $\text{SU}(2)$ fields of periods two, four and six gives $h_F = 13.97401$, attained by an abelian field; certificates made of plaquette forms and vertex forms on the period-two cell cannot go below 14.384. With this gauge the curvature method stops between $|\beta| = \frac{1}{29}$ and $\frac{1}{26.25}$.

VII. SU(3) UP TO 3/4

In this section $G = \text{SU}(3)$, $\beta \geq 0$ and $\beta_W = 9\beta$. The distance on $\text{SU}(3)$ is $d(U, V) = \|U - V\|_{\text{HS}}$, which is at most the geodesic distance for $\langle X, Y \rangle = -\text{Tr}(XY)$; $K = \text{conv } \text{SU}(3)$; $A(M) = \frac{1}{2}(M - M^*)$ and P removes the trace. With this inner product $\text{Ric} = \frac{3}{2}$, the Casimir of the fundamental representation is $\lambda_F = \frac{8}{3}$ and that of the adjoint representation is 6.

A. Reduction to one link

The conditional law of an active link is

$$\nu_b(dU) \propto e^{\text{Re Tr}(U b)} dU, \quad b = N\beta \sum_{p \ni e} S_p,$$

where S_p is the staple of e in p , a product of $\text{SU}(3)$ matrices and identities, or its adjoint. Six staples give $b \in rK$ with $r = 6N\beta = 2\beta_W$. For a matrix h put $\varphi_h(U) = \text{Re Tr}(Uh)$ and

$$L_h(b) = \sup\{\text{Cov}_{\nu_b}(g, \varphi_h) : g \text{ 1-Lipschitz for } d\},$$

$$\Lambda(r) = \sup_{b \in rK} \sup_{\|h\|_{\text{HS}}=1} L_h(b).$$

Changing a partner e' changes one staple by a matrix of norm $d(Q_{e'}, Q'_{e'})$. Along the segment b_t from b to b' , which stays in rK , $\frac{d}{dt} \nu_{b_t}(g) = \text{Cov}_{\nu_{b_t}}(g, \text{Re Tr}(U(b' - b)))$, so $W_1(\nu_b, \nu_{b'}) \leq N\beta \Lambda(r) d(Q_{e'}, Q'_{e'})$. With twelve partners per link,

$$\alpha \leq 12N\beta \Lambda(r) = 2r\Lambda(r) = 4\beta_W \Lambda(2\beta_W). \quad (8)$$

Since K is convex and contains 0, the mean of $I, \omega I, \omega^2 I$ with $\omega^3 = 1$, $r'K \subset rK$ for $r' \leq r$ and Λ is nondecreasing.

B. Gradients and one-forms

Lemma 27. *In the left trivialisation, $\nabla \varphi_c(U) = -PA(cU)$; if $V = \varphi_b$ then $\text{Hess } V(X, X) = \text{Re Tr}(UX^2 b)$; and $\Delta \varphi_c = -\lambda_F \varphi_c$.*

Proof. $\frac{d}{dt} \text{ReTr}(Ue^{tX}c)|_{t=0} = \text{ReTr}(cUX) = -\langle PA(cU), X \rangle$ for $X \in \mathfrak{su}(3)$. On a group with a bi-invariant metric $\nabla_X X = 0$ for left-invariant X , so the Hessian is the second derivative along e^{tX} . The function φ_c is a matrix coefficient of the fundamental representation, on which the Casimir acts by λ_F . $\square$

Lemma 28. *The bottom of the spectrum of the rough Laplacian $\nabla^* \nabla$ on L^2 one-forms of $\text{SU}(3)$ is $\mu_0 = \frac{7}{6}$.*

Proof. Trivialise one-forms on the left, so that a one-form is a function $f : \text{SU}(3) \rightarrow \mathfrak{su}(3)$. For an orthonormal basis X_a acting as left-invariant vector fields, $\nabla_{X_a} Y = \frac{1}{2} [X_a, Y]$ and $\nabla_{X_a} X_a = 0$, so $\nabla^* \nabla = -\sum_a (\rho(X_a) + \frac{1}{2} \text{ad} X_a)^2$. On the isotypic part $V_\pi \otimes V_\pi^* \otimes \mathfrak{su}(3)$ of $L^2 \otimes \mathfrak{su}(3)$ put $A_a = \rho_\pi(X_a)$, $B_a = \text{ad} X_a$ and $T_a = A_a + B_a$, the action on the tensor product. Since A_a and B_a commute, $A_a + \frac{1}{2} B_a = \frac{1}{2} (A_a + T_a)$ and $A_a T_a + T_a A_a = A_a^2 + T_a^2 - B_a^2$, so on a component $\sigma \subset \pi \otimes \text{ad}$

$$\nabla^* \nabla = \frac{1}{2} (C_\pi + C_\sigma) - \frac{1}{4} C_{\text{ad}} = \frac{1}{2} (C_\pi + C_\sigma) - \frac{3}{2},$$

with $C(p, q) = \frac{2}{3} (p^2 + q^2 + pq + 3p + 3q)$ for the representation of highest weight (p, q) [34]. The components of $(p, q) \otimes (1, 1)$ are $(p+1, q+1)$, $(p+2, q-1)$, $(p-1, q+2)$, $(p+1, q-2)$, $(p-2, q+1)$, $(p-1, q-1)$ and (p, q) , the last with multiplicity two, one or zero as both, one or none of p, q are positive, and only those with nonnegative labels occur. For $\pi = (0, 0)$ the only component is $(1, 1)$, with value $\frac{3}{2}$. For $\pi = (1, 0)$ the components are $(2, 1)$, $(0, 2)$, $(1, 0)$, with values $\frac{31}{6}$, $\frac{19}{6}$, $\frac{7}{6}$, and $(0, 1)$ is conjugate. For $p+q \geq 2$ the smallest values occur at $(p-1, q-1)$, (p, q) , $(p-2, q+1)$ or $(p+1, q-2)$, and each is at least $\frac{1}{3} (2(p^2 + q^2 + pq) + 3(p+q) - 3) - \frac{3}{2} \geq \frac{3}{2}$. $\square$

C. The gap of the one-link measure

Lemma 29. *For $b \in rK$ the spectral gap of $L = \Delta + \langle \nabla V, \nabla \cdot \rangle$, $V = \varphi_b$, on $L^2(\nu_b)$ is at least*

$$\gamma(r) = \frac{39}{64} \mu_0 + \frac{3}{2} - \frac{r}{2} = \frac{283}{128} - \frac{r}{2}.$$

Proof. Let $-Lf = \lambda f$ with $\lambda > 0$, $\omega = df$ and $\eta = e^{V/2} \omega$; norms are in $L^2(dU)$. Bochner's formula integrated against ν_b gives $\lambda \nu_b |\omega|^2 = \nu_b |\nabla \omega|^2 + \nu_b ((\text{Ric} - \text{Hess } V)(\omega, \omega))$, and since $e^{V/2} \nabla \omega = \nabla \eta - \frac{1}{2} dV \otimes \eta$,

$$\lambda \|\eta\|^2 = \|\nabla \eta - \frac{1}{2} dV \otimes \eta\|^2 + \int (\text{Ric} - \text{Hess } V)(\eta, \eta). \quad (9)$$

Put $\mathcal{A} = \nabla \eta$ and $\mathcal{B} = \frac{1}{2} dV \otimes \eta$. Then

$$\langle \mathcal{A}, \mathcal{B} \rangle = \frac{1}{4} \int \langle \nabla V, \nabla |\eta|^2 \rangle = \frac{\lambda_F}{4} \int V |\eta|^2,$$

because $-\Delta V = \lambda_F V$. Writing $\|\mathcal{A} - \mathcal{B}\|^2 = \|\frac{5}{8} \mathcal{A} - \mathcal{B}\|^2 + \frac{39}{64} \|\mathcal{A}\|^2 - \frac{3}{4} \langle \mathcal{A}, \mathcal{B} \rangle$ and dropping the first term,

$$\|\mathcal{A} - \mathcal{B}\|^2 \geq \frac{39}{64} \mu_0 \|\eta\|^2 - \frac{1}{2} \int V |\eta|^2$$

by Lemma 28. With $\text{Ric} = \frac{3}{2}$, (9) gives $\lambda \|\eta\|^2 \geq (\frac{39}{64} \mu_0 + \frac{3}{2}) \|\eta\|^2 - \int (\text{Hess } V(\eta, \eta) + \frac{1}{2} V |\eta|^2)$, and the claim follows from the pointwise bound

$$\text{ReTr}(MX^2) + \frac{1}{2} \text{ReTr} M \leq \frac{r}{2} |X|^2, \quad M = bU \in rK, \quad (10)$$

since $\text{Hess } V(X, X) = \text{ReTr}(MX^2)$ and $V = \text{ReTr} M$. The maximum over $|X| = 1$ of the left side is convex in M , so it suffices to take $M = rW$ with $W \in \text{SU}(3)$ and, after a conjugation, $W = \text{diag}(e^{i\theta_k})$ with $\sum_k \theta_k = 0$; put $c_k = \cos \theta_k$. Write $X = iH$ with H Hermitian, traceless, $\text{Tr} H^2 = 1$, with diagonal d_k and entries H_{jk} . Then

$$\text{ReTr}(WX^2) = -\sum_k c_k d_k^2 - \sum_{j < k} (c_j + c_k) |H_{jk}|^2,$$

with $\sum_k d_k^2 + 2 \sum_{j < k} |H_{jk}|^2 = 1$, so the maximum is the larger of $q_{\text{off}} = \max_{j < k} -\frac{1}{2} (c_j + c_k)$ and $q_{\text{diag}} = \max \{ -\sum_k c_k d_k^2 : \sum d_k = 0, \sum d_k^2 = 1 \}$. If l is the third index, $q_{\text{off}} + \frac{1}{2} \sum_k c_k = \frac{1}{2} c_l \leq \frac{1}{2}$. The value q_{diag} is minus the smaller eigenvalue of $\text{diag}(c)$ compressed to the plane $\sum d_k = 0$; the eigenvalues of the compression solve $\sum_k (c_k - \mu)^{-1} = 0$, that is $3\mu^2 - 2e_1\mu + e_2 = 0$ with e_1, e_2 the elementary symmetric functions of the c_k , so $q_{\text{diag}} = \frac{1}{3} (\sqrt{e_1^2 - 3e_2} - e_1)$. The inequality $q_{\text{diag}} + \frac{1}{2} e_1 \leq \frac{1}{2}$ is $\sqrt{e_1^2 - 3e_2} \leq \frac{1}{2} (3 - e_1)$, that is $e_1^2 - 4e_2 + 2e_1 - 3 \leq 0$. When $\sum \theta_k = 0$, $\sum c_k^2 = 1 + 2c_1 c_2 c_3$, and with this identity the left side equals $-2(1 - c_1)(1 - c_2)(1 - c_3) \leq 0$. $\square$

The weight $\frac{5}{8}$ is the one that gives the coefficient $\frac{1}{2}$ of V , and (10) is an equality at $W = I$. At $r = \frac{3}{2}$, $\gamma = \frac{187}{128}$.¹²

D. A corrector of degree two

Lemma 30. *For $b \in rK$, $\|h\| = 1$ and every smooth real function u on $\text{SU}(3)$,*

$$L_h(b) \leq \sqrt{\nu_b |\nabla u|^2} + \sqrt{\text{Var}_{\nu_b}(\varphi_h + Lu) / \gamma(r)}.$$

Proof. Let $-L\psi = \varphi_h - \nu_b \varphi_h$. For a 1-Lipschitz g , $|\nabla g| \leq 1$ almost everywhere, because d is at most the geodesic distance, so $\text{Cov}(g, \varphi_h) = \nu_b \langle \nabla g, \nabla \psi \rangle \leq \nu_b |\nabla \psi|$. Write $\psi = u + \chi$. Since $\nu_b(Lu) = 0$, χ solves $-L\chi = X - \nu_b X$ with $X = \varphi_h + Lu$, and

$$\nu_b |\nabla \chi|^2 = \text{Cov}(\chi, X) \leq \sqrt{\text{Var } \chi \text{Var } X} \leq \sqrt{\nu_b |\nabla \chi|^2 \text{Var } X / \gamma}$$

by Lemma 29. Hence $\nu_b |\nabla \chi|^2 \leq \text{Var } X / \gamma$, and $\nu_b |\nabla \psi| \leq \sqrt{\nu_b |\nabla u|^2} + \sqrt{\nu_b |\nabla \chi|^2}$. $\square$

¹² The direct Bochner bound, with Minkowski's inequality in place of the cross term, gives $(\sqrt{\mu_0} - \frac{5}{8}r)^2 + \frac{3}{2} - r$ with $g^2 = \max_W \|PA(W)\|^2 = 2.0655$; it is exact at $b = 0$, where the gap is $\frac{8}{3}$, but falls below γ for $r > 0.24$ and vanishes before $r = 1$.

The best u is ψ itself, and then the second term vanishes. With $u = s\varphi_h$, Lemma 27 gives $\varphi_h + Lu = (1 - s\lambda_F)\varphi_h + s\Gamma$ with $\Gamma = \langle \nabla V, \nabla \varphi_h \rangle$; the choice $s = 1/\lambda_F$ removes φ_h and leaves the drift. This blended corrector reaches $\beta_W = \frac{45}{64}$ and stops near 0.715 (Sec. IX). We take u in a space of polynomials of degree two.

Let $\mathcal{P}_2$ be the real vector space of real polynomials of degree one or two in the entries u_{ij} of U and $\bar{u}_{ij}$ of $\bar{U}$. It has dimension 189: a basis (B_k) consists of $\frac{1}{2}(m + \bar{m})$ and $\frac{1}{2i}(m - \bar{m})$ for each pair $\{m, \bar{m}\}$ of conjugate monomials with $m \neq \bar{m}$, and m itself when $m = \bar{m}$. Identify $h \in \mathbb{C}^{3 \times 3}$ with $\mathbb{R}^{18}$ through the real and imaginary parts of its entries. For a real 189×18 matrix C put

$$q_h = \sum_k \sum_{m=1}^{18} C_{km} h_m B_k \in \mathcal{P}_2,$$

linear in h . For diagonal $b = \text{diag}(a_1, a_2, x_3 + iy_3)$ we have $V = \sum_k b_k \varphi_{F_k}$ with $(b_k) = (a_1, a_2, x_3, y_3)$ and $F = (e_{11}, e_{22}, e_{33}, ie_{33})$, so

$$X_h = \varphi_h + Lq_h = \varphi_h + \Delta q_h + \sum_k b_k \langle \nabla \varphi_{F_k}, \nabla q_h \rangle \quad (11)$$

is a polynomial of degree at most three, affine in b , while ∇q_h does not depend on b . The squares of these functions have degree at most six, so the quadratic forms $h \mapsto v_b |\nabla q_h|^2$ and $h \mapsto v_b [(X_h - \kappa)^2]$ are combinations of moments of v_b of degree at most six.

To compute Δ and $\langle \nabla f, \nabla g \rangle$ exactly we use the basis of $\mathfrak{su}(3)$

$$\begin{aligned} e_{jk} - e_{kj}, \quad i(e_{jk} + e_{kj}) \quad (j < k), \\ i(e_{11} - e_{22}), \quad i(e_{11} + e_{22} - 2e_{33}), \end{aligned}$$

acting as left-invariant fields $X_a f(U) = \frac{d}{dt} f(U e^{tX_a})|_{t=0}$, with weights $w_a = \frac{1}{2}$ for the first seven and $\frac{1}{6}$ for the last. Then $\sum_a w_a X_a \otimes X_a$ is the Casimir tensor of $\langle X, Y \rangle = -\text{Tr}(XY)$, so

$$\langle \nabla f, \nabla g \rangle = \sum_a w_a X_a f X_a g, \quad \Delta = \sum_a w_a X_a^2.$$

The fields map monomials to combinations of monomials of the same degree with coefficients in $\mathbb{Z}[i]$; for instance $\Delta \varphi_h = -\frac{8}{3} \varphi_h$, as in Lemma 27.

E. Orbits of staple sums

Lemma 31. *Let $V, W \in \text{SU}(3)$.*

- (i) $L_h(VbW) = L_{V^{-1}hW^{-1}}(b)$ and $L_h(\bar{b}) = L_{\bar{h}}(b)$.
- (ii) Every $b \in \mathbb{C}^{3 \times 3}$ is in the orbit of some $\text{diag}(a_1, a_2, z_3)$ with $a_1 \geq a_2 \geq |z_3|$ and $\text{Im } z_3 \geq 0$.
- (iii) A diagonal matrix lies in K if and only if it lies in the convex hull of the diagonal torus T of $\text{SU}(3)$.

Proof. (i) If U has law v_{VbW} then WUV has law v_b ; the map preserves d and Haar measure, and $\varphi_h(U) = \varphi_{V^{-1}hW^{-1}}(WUV)$.

For $\bar{b}$ use $U \mapsto \bar{U}$. (ii) Take a singular value decomposition $b = V_1 \Sigma W_1$ and move the determinants of V_1, W_1 into the last entry; a complex conjugation makes $\text{Im } z_3 \geq 0$. (iii) Let c be diagonal with nonzero entries of distinct moduli and let U maximise $\text{Re Tr}(cU)$ on $\text{SU}(3)$. The derivative along $\mathfrak{su}(3)$ vanishes, so the anti-Hermitian part of cU is a multiple $i\lambda I$ of the identity and $cU = H + i\lambda I$ with H Hermitian. Then $H^2 + \lambda^2 I = cc^*$ is diagonal with distinct entries, H commutes with it and is diagonal, and $U \in T$. Hence $\max_{\text{SU}(3)} \text{Re Tr}(cU) = \max_T \text{Re Tr}(ct)$ for such c , and by continuity for every diagonal c . A diagonal matrix outside $\text{conv } T$ is separated from it by such a functional, which then separates it from K . $\square$

Since every $b \in rK$ has operator norm at most r ,

$$\Lambda(r) = \sup \left\{ \sup_{\|h\|=1} L_h(b) : b \in \mathcal{R}_r \right\},$$

where $\mathcal{R}_r$ is the set of $\text{diag}(a_1, a_2, x_3 + iy_3) \in r \text{conv } T$ with $a_1 \geq a_2 \geq |x_3 + iy_3|$ and $y_3 \geq 0$, a subset of the box $[0, r]^2 \times [-r, r] \times [0, r]$ in the coordinates (a_1, a_2, x_3, y_3) .

F. Exact moments and interpolation

Lemma 32. *For $D, D' \in \mathbb{C}^{3 \times 3}$ let $\mathbb{E}_H$ be the Haar mean on $\text{SU}(3)$ and*

$$Z(D, D') = \mathbb{E}_H \exp\left(\frac{1}{2} \text{Tr}(UD) + \frac{1}{2} \text{Tr}(\bar{U}D')\right).$$

Then

$$Z(D, D') = \sum_{p \geq q \geq 0} \frac{s_{(p,q)}(DD^\top) G_{p,q}(\det D) G_{p,q}(\det D')}{d_{p,q}},$$

with $G_{p,q}(z) = \sum_{k \geq 0} c_{(p+q+k, k)} z^k$, $c_\lambda = 2^{-|\lambda|} / \prod \text{hooks}(\lambda)$, $d_{p,q} = \frac{1}{2}(p-q+1)(q+1)(p+2)$ and $s_{(p,q)}$ the Schur polynomial. The moments of degree j in U and k in $\bar{U}$ are $\mathbb{E}_H[u^\alpha \bar{u}^\beta e^{\text{Re Tr}(Ub)}] = 2^{|\alpha|+|\beta|} \partial_D^\alpha \partial_{D'}^\beta Z(b, \bar{b})$, the entry u_{ij} pairing with D_{ji} .

Proof. We use $(\text{Tr } X)^n = \sum_{\lambda \vdash n} f^\lambda s_\lambda(X)$, where f^λ is the number of standard tableaux of shape λ , given by the hook length formula [71, Cor. 7.21.6]. Expanding the exponential gives

$$e^{\frac{1}{2} \text{Tr } X} = \sum_{n \geq 0} \frac{(\text{Tr } X)^n}{2^n n!} = \sum_{\lambda} c_\lambda s_\lambda(X),$$

the sum over partitions with at most three parts. For $U \in \text{SU}(3)$ one has $s_{\lambda+(k,k,k)}(UD) = (\det D)^k s_\lambda(UD)$, so

$$e^{\frac{1}{2} \text{Tr}(UD)} = \sum_{p \geq q} G_{p,q}(\det D) \chi_{p,q}(UD),$$

where $\chi_{p,q}$ is the irreducible character of highest weight $(p, q, 0)$ and $d_{p,q}$ its dimension. Since $\bar{U} = (U^{-1})^\top$, we have $\chi_{p,q}(\bar{U}D') = \chi_{p,q}(U^{-1}D'^\top)$, and Schur orthogonality [22, Ch. 2],

$$\mathbb{E}_H[\chi_p(UA) \chi_{p'}(U^{-1}B)] = \delta_{pp'} \frac{\chi_p(AB)}{d_p},$$

gives the expansion. The bound below gives absolute convergence. Finally $\text{Re Tr}(Ub) = \frac{1}{2} \text{Tr}(Ub) + \frac{1}{2} \text{Tr}(\bar{U}\bar{b})$. $\square$

The computation keeps $p \leq 34$ and $k \leq 11$. Every omitted partition has more than 34 boxes, and since $|s_\lambda(X)| \leq \dim(\lambda) \|X\|^{|\lambda|}$ and $\sum_{\lambda \vdash n} f^\lambda \dim \lambda = 3^n$ by Schur–Weyl duality, the omitted part of each factor is at most $\sum_{n \geq 34} (\frac{3}{2} \|D\|)^n / n!$; Cauchy’s estimate on a polydisc around b turns this into a bound for every derivative (Appendix E).

Lemma 33 (Interpolation on a box). *Let $Q = \prod_k [l_k, l_k + w_k]$ in the coordinates (a_1, a_2, x_3, y_3) , with corners b_i and multi-linear weights $\omega_i(b) \geq 0$, $\sum_i \omega_i(b) = 1$, and let $\varsigma = w_1^2 + w_2^2 + \max(w_3^2, w_4^2)$. For $b \in Q$:*

- (i) $Z(b) \geq e^{-\varsigma/8} \sum_i \omega_i(b) Z(b_i)$;
- (ii) $v_b(f) \leq e^{\varsigma/8} \max_i v_{b_i}(f)$ for every $f \geq 0$ that does not depend on b ;
- (iii) if $Y_b = (\mathcal{A} + \sum_k (b_k - \beta_k) \mathcal{B}_k)^2$ with functions $\mathcal{A}, \mathcal{B}_k$ of U and a fixed point β , then

$$v_b(Y_b) \leq e^{\varsigma/8} \left(\max_i v_{b_i}(Y_{b_i}) + \sum_k \frac{w_k^2}{4} \frac{\max_i Z(b_i) v_{b_i}(\mathcal{B}_k^2)}{\min_i Z(b_i)} \right).$$

Proof. Write $E(b) = e^{\text{Re Tr}(Ub)} = \prod_k e^{c_k b_k}$, where

$$c = (\text{Re } u_{11}, \text{Re } u_{22}, \text{Re } u_{33}, -\text{Im } u_{33}),$$

so that $|c_k| \leq 1$ and $c_3^2 + c_4^2 \leq 1$. In each coordinate the chord of $e^{c_k x}$ over $[l_k, l_k + w_k]$ lies above the function and, by Hoeffding’s lemma [45] for a two-point variable, below $e^{c_k x} e^{c_k^2 w_k^2 / 8}$. Multiplying,

$$E(b) \leq \sum_i \omega_i(b) E(b_i) \leq E(b) e^{\sum_k c_k^2 w_k^2 / 8} \leq E(b) e^{\varsigma/8}. \quad (12)$$

Integrating the right inequality gives (i), and (ii) follows from the left one and (i). For (iii) let $\mathcal{F}(b) = Y_b E(b)$ and I_k the linear interpolation in the k -th coordinate, so that $\mathcal{F} - I_1 I_2 I_3 I_4 \mathcal{F} = \sum_k I_1 \cdots I_{k-1} (\mathcal{F} - I_k \mathcal{F})$. Along the k -th coordinate $\mathcal{F}$ is $\ell^2 e^{c_k x}$ times a positive factor, with ℓ affine of slope $\mathcal{B}_k$, and $\partial_k^2 \mathcal{F} = ((c_k \ell + 2\mathcal{B}_k)^2 - 2\mathcal{B}_k^2) E \geq -2\mathcal{B}_k^2 E$. With the Green kernel of the interval, whose integral is at most $w_k^2 / 8$,

$$\mathcal{F} - I_k \mathcal{F} \leq 2\mathcal{B}_k^2 \int G E \leq \frac{w_k^2}{4} \mathcal{B}_k^2 \sum_i \omega_i^{(k)} E(b_i)$$

by the left inequality of (12), with weights $\omega_i^{(k)} \geq 0$ of sum one that do not depend on U . The operators I_j have nonnegative weights of sum one; integrate over U and divide by $Z(b)$, bounded below by (i). $\square$

G. The cover and the theorem

Fix a box Q and a matrix C . For any constant κ , $\text{Var}(X_h) \leq v_b[(X_h - \kappa)^2]$; take $\kappa = \langle \kappa(b), h \rangle$ with $\kappa(b)$ affine in b , fitted to the means at the corners. By (11), $X_h - \kappa = \mathcal{A} +$

$\sum_k (b_k - \beta_k) \mathcal{B}_k$ with β the centre of Q and $\mathcal{B}_k = \langle \nabla \phi_{F_k}, \nabla q_h \rangle - \langle \partial_k \kappa, h \rangle$, and Lemmas 30 and 33 give, for all $b \in Q \cap \mathcal{R}_r$,

$$\sup_h L_h(b) \leq \sqrt{e^{\varsigma/8} \max_i \lambda_A(b_i)} + \sqrt{\frac{e^{\varsigma/8}}{\gamma(r)} \left(\max_i \lambda_V(b_i) + \frac{e_N}{\min_i Z(b_i)} \right)}, \quad (13)$$

where $e_N = \sum_k \frac{w_k^2}{4} \max_i Z(b_i) \lambda_k(b_i)$ and $\lambda_A, \lambda_V, \lambda_k$ are the largest eigenvalues of the quadratic forms $h \mapsto v_{b_i} |\nabla q_h|^2$, $h \mapsto v_{b_i} [(X_h - \kappa)^2]$ at $b = b_i$, and $h \mapsto v_{b_i} [\mathcal{B}_k^2]$ on $\mathbb{R}^{18}$. Their entries are moments of v_{b_i} of degree at most six, computed with rigorous error bounds (Appendix E); each largest eigenvalue is bounded above by an $LDL^\top$ factorisation of tI minus the computed form with certified positive pivots, plus the norm of its error matrix by Weyl’s inequality [43, 48].

The matrix C is chosen in each box in floating point. At the mean of the corner moments and at $b = \beta$, Lemma 30 for $u = q_h$ is a quadratic problem in C : for $\theta > 0$ we minimise the trace of $v |\nabla q|^2 + \theta \text{Var}(X) / \gamma$ over the 18 coordinate directions h , a least-squares problem of size 189, keep the θ on a geometric grid of 48 values from 0.05 to 80 that gives the smallest bound, and round C to multiples of 2^{-30} . Any choice gives a valid bound.

A box is discarded when it contains no point with $a_1 \geq a_2 \geq |z_3|$, or when a linear functional $\ell(b) = \text{Re Tr}(c^* b)$ satisfies $\min_Q \ell > r \max_T \ell$, which separates it from $r \text{conv } T$ and hence, by Lemma 31(iii), from $\mathcal{R}_r$.

Proposition 34. $2r\Lambda(r) \leq 0.999887$ at $r = \frac{3}{2}$.

Proof. The box $[0, r]^2 \times [-r, r] \times [0, r]$ is cut into 8192 boxes of side $r/8$. Of these, 841 survive the two tests, and 581 satisfy $2rB_Q < 1$ for the right side B_Q of (13), with $\gamma(r) = \frac{187}{128}$. Each of the other 260 is cut into sixteen boxes of side $r/16$; 1760 of these survive the tests, and all satisfy $2rB_Q < 1$. The 2341 certified boxes cover $\mathcal{R}_r$. The largest bound is 0.999887, on the box $[1.3125, 1.5] \times [0.9375, 1.125] \times [0.5625, 0.75] \times [0, 0.1875]$, and the median is 0.983. Figure 6 shows the cover, and Appendix E the error analysis. $\square$

The same cover at a larger r bounds $\Lambda(r)$ where $2r\Lambda(r)$ is no longer below one; the plaquette blocks of Section VIII use that bound.

Proposition 35. $2r\Lambda(r) \leq 1.029967$ at $r = \frac{38}{25}$.

Proof. As in the proof of Proposition 34, with $\gamma(r) = \frac{4643}{3200}$ and the threshold $2rB_Q < 1.03$: of the 8192 boxes of side $r/8$, 784 survive the two tests and 690 satisfy the threshold; each of the other 94 is cut into sixteen boxes of side $r/16$, of which 373 survive the tests, and all satisfy it. The 1063 certified boxes cover $\mathcal{R}_r$, and the largest bound is 1.029967. $\square$

Proof of Theorem 6. Proposition 34 gives

$$\Lambda\left(\frac{3}{2}\right) \leq \frac{0.999887}{3} < 0.333296.$$

For $\beta_W \leq \frac{3}{4}$ we have $2\beta_W \leq \frac{3}{2}$, and (8) with the monotonicity of Λ gives

$$\alpha \leq 4\beta_W \Lambda\left(\frac{3}{2}\right) < 1.333183\beta_W \leq \frac{3}{4} \cdot 1.333183 < 1.$$

Proposition 11 applies with the Hilbert–Schmidt metric. $\square$

VIII. SU(3) WITH PLAQUETTE BLOCKS

In this section $G = \text{SU}(3)$, $\beta \geq 0$, $\beta_W = 9\beta$, $N\beta = \beta_W/3$, and the notation is that of Sec. VII: d is the Hilbert–Schmidt distance, $K = \text{conv SU}(3)$, $v_c(dU) \propto e^{\text{Re Tr}(Uc)} dU$, $\varphi_h(U) = \text{Re Tr}(Uh)$ and $\Lambda(r)$ is the modulus of (8). Put $Z(c) = \int e^{\text{Re Tr}(Uc)} dU$ and $g = \log Z$ on $\mathbb{C}^{3 \times 3}$, viewed as $\mathbb{R}^{18}$ with the inner product $\langle x, y \rangle = \text{Re Tr}(xy^*)$. The gradient of g at c is the conjugate mean $m(c)^*$, $m(c) = v_c(U) \in K$, and $\text{Hess } g(c)(h, h) = \text{Var}_{v_c}(\varphi_h)$. Since $Z(VcW) = Z(c)$ for $V, W \in \text{SU}(3)$, the same holds for g .

The blocks are those of Lemma 12, which concerns the line field only and holds for every group. For a block plaquette with active links e_1, e_2, e_3 , $\text{Re Tr } Q_p = \text{Re Tr}(y_1 y_2 y_3)$ with $y_i = Q_{e_i}$ or $Q_{e_i}^{-1}$, and the conditional law of the block is

$$\mu_b(dy) \propto \exp\left(\sum_i \text{Re Tr}(y_i b_i) + N\beta \text{Re Tr}(y_1 y_2 y_3)\right) dy, \quad (14)$$

where b_i is $N\beta$ times the sum of the five staples of y_i outside the block plaquette, so $b_i \in 5N\beta K$ and $\|b_i\|_{\text{op}} \leq 5N\beta$. On $\text{SU}(3)^3$ we use $d(y, y') = \sum_i \|y_i - y'_i\|_{\text{HS}}$.

A. The sensitivity of one block

Let $r = 2\beta_W = 6N\beta$. Let λ_* and τ be a lower and an upper bound for the eigenvalues of $\text{Hess } g$ on rK , M an upper bound on rK for $|D^3 g(c)(u, v, v)|$ over unit vectors u, v , and $\gamma = \gamma(r)$ the gap of Lemma 29. Put

$$\begin{aligned} w &= \frac{3}{2}(N\beta)^2(\tau - \lambda_*), & s &= \Lambda(r) + (e^w - 1)e^w \sqrt{\tau/\gamma}, \\ C &= 5(N\beta)^2 \left(\tau s + \frac{\sqrt{3}N\beta M e^w}{\gamma} \right), \\ K_3 &= \frac{(1 + N\beta\Lambda(r))s}{1 - C}, & \alpha_3 &= 10N\beta K_3, \end{aligned} \quad (15)$$

assuming $C < 1$.

Lemma 36. *Let $a, c_0 \in 5N\beta K$ and*

$$\kappa(dx) \propto \exp(\text{Re Tr}(xa) + g(c_0 + N\beta x)) dx \quad \text{on } \text{SU}(3).$$

For every 1-Lipschitz f and every $h \in \mathbb{C}^{3 \times 3}$:

- (i) $\text{Cov}_\kappa(f, \varphi_h) \leq s\|h\|_{\text{HS}}$;
- (ii) *if c_0 moves in the direction k inside $5N\beta K$, then*

$$\left| \frac{d}{dt} \kappa(f) \right| \leq N\beta \left(\tau s + \sqrt{3}N\beta M e^w / \gamma \right) \|k\|_{\text{HS}}.$$

Proof. (i) The segment from c_0 to $c_0 + N\beta x$ lies in $6N\beta K = rK$, and Taylor's formula gives

$$g(c_0 + N\beta x) = g(c_0) + N\beta \text{Re Tr}(x m(c_0)) + R(x)$$

with $R(x) = \frac{1}{2}(N\beta)^2 \text{Hess } g(\xi)(x, x)$. Since $\|x\|_{\text{HS}}^2 = 3$, $R \in [\frac{3}{2}(N\beta)^2 \lambda_*, \frac{3}{2}(N\beta)^2 \tau]$ and $\text{osc } R \leq w$. Hence κ is the tilt $e^R v / v(e^R)$ of $v = v_{c'}$ with $c' = a + N\beta m(c_0) \in rK$. By definition of Λ , $\text{Cov}_v(f, \varphi_h) \leq \Lambda(r)\|h\|$. The function f is 1-Lipschitz for the geodesic distance, which dominates d , so $\text{sd}_v f \leq \gamma^{-1/2}$ by Lemma 29, and $\text{sd}_v \varphi_h \leq \sqrt{\tau}\|h\|$. Lemma 13 gives (i).

(ii) The derivative is $\text{Cov}_\kappa(f, \langle \nabla g(c_0 + N\beta x), k \rangle)$. Write $\nabla g(c_0 + N\beta u) = \nabla g(c_0) + N\beta \text{Hess } g(c_0)u + N(u)$. The linear term is φ_h with $\|h\| \leq N\beta \tau \|k\|$ and contributes at most $N\beta \tau s \|k\|$ by (i). On the segment from c_0 the Hessian moves by at most $N\beta M \|u\| \leq \sqrt{3}N\beta M$, so $x \mapsto \langle \nabla g(x), k \rangle$ is $\sqrt{3}(N\beta)^2 M \|k\|$ -Lipschitz for d . The law κ is a tilt of v by a function of oscillation at most w , so its spectral gap is at least $e^{-w}\gamma$ by the Holley–Stroock lemma [44], and the covariance of two Lipschitz functions under κ is at most e^w/γ times the product of their constants. $\square$

Lemma 37. *Let $C < 1$. If b, b' have all fields in $5N\beta K$ and differ only in b_i , then $W_1(\mu_b, \mu_{b'}) \leq K_3 \|b_i - b'_i\|_{\text{HS}}$ for the metric d .*

Proof. By the cyclic symmetry of $\text{Re Tr}(y_1 y_2 y_3)$ we may take $i = 1$. Given y_1, y_3 , the law of y_2 is v_c with $c = b_2 + N\beta y_3 y_1 \in rK$. Changing y_1 by η in d moves c by $N\beta \eta$ and v_c by at most $N\beta \Lambda(r)\eta$ in W_1 , as in (8); so $F(y_1, y_3) = \mathbb{E}[f | y_1, y_3]$ is $(1 + N\beta \Lambda(r))$ -Lipschitz in each variable. The law of (y_1, y_3) is proportional to $e^{\text{Re Tr}(y_1 b_1) + \text{Re Tr}(y_3 b_3)} Z(b_2 + N\beta y_3 y_1)$, and $Z(b_2 + N\beta y_3 y_1) = Z(y_3^{-1} b_2 + N\beta y_1) = Z(b_2 y_1^{-1} + N\beta y_3)$, so its two conditional laws have the form of Lemma 36, with $c_0 = y_3^{-1} b_2$ and $c_0 = b_2 y_1^{-1}$. Moving y_3 to y'_3 moves $y_3^{-1} b_2$ along a segment in $5N\beta K$ of length at most $\|b_2\|_{\text{op}} \|y_3 - y'_3\| \leq 5N\beta \|y_3 - y'_3\|$. The coupling argument of Lemma 15 now gives $d_1 \leq s\varepsilon + Cd_3$, $d_3 \leq Cd_1$ and $|\mu_b(f) - \mu_{b'}(f)| \leq (1 + N\beta \Lambda(r))(d_1 + d_3) \leq K_3 \varepsilon$. $\square$

B. The constants and the proof of Theorem 7

Lemma 38. *At $r = \frac{3}{2}$ every eigenvalue of $\text{Hess } g$ on rK lies in $[0.056015, 0.247316]$, and $|D^3 g(c)(u, v, v)| \leq 0.182126$ for $c \in rK$ and unit vectors u, v . At $r = \frac{38}{25}$ the same holds with $[0.054978, 0.248280]$ and 0.183546 .*

Proof. For $V, W \in \text{SU}(3)$, $g(VcW) = g(c)$ and $x \mapsto VxW$ is an isometry, so the eigenvalues of $\text{Hess } g$ and the norm of $D^3 g$ are constant on the orbits, and by Lemma 31 it suffices to bound them on $\mathcal{R}$. The derivatives of g are the cumulants of φ under v_c . Cut the box $[0, r]^2 \times [-r, r] \times [0, r]$ into boxes of side $r/8$ and keep the 841 (784 at $r = \frac{38}{25}$) that survive the two tests of Sec. VII G. On each box Q we compute, with rigorous error bounds, the moments of φ of degree at most three at the sixteen corners b_i and at the centre b_0 , and put $k = m(b_0)$, the mean of φ at the centre.

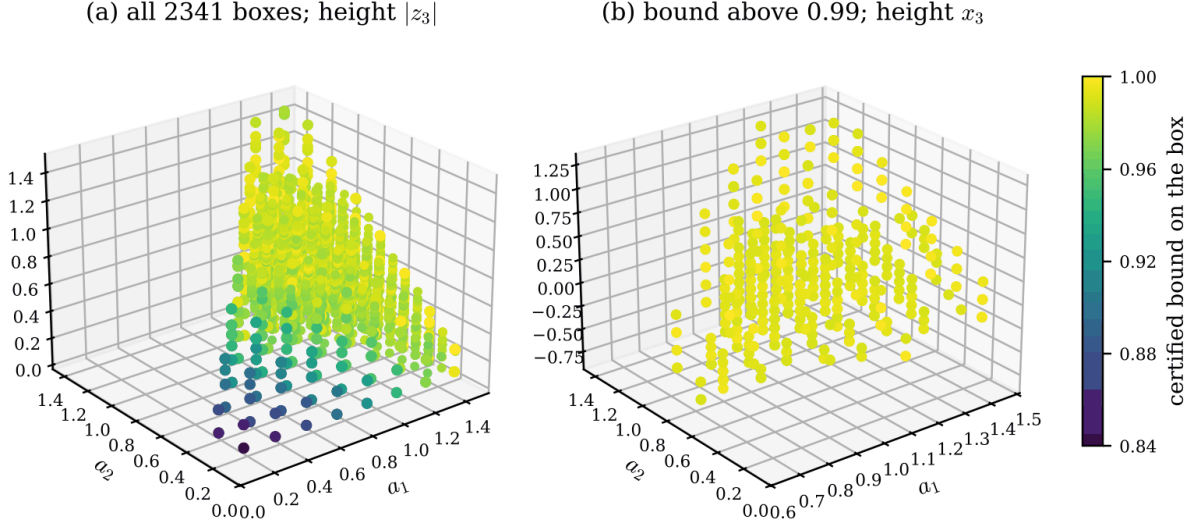


FIG. 6. The 2341 certified boxes at $r = \frac{3}{2}$, plotted at their centres in the coordinates (a_1, a_2) of the staple sum $\text{diag}(a_1, a_2, x_3 + iy_3)$, coloured by the certified bound on $2r \sup_h L_h$ over the box; the larger points are the boxes of side $r/8$. (a) All boxes, with height $|x_3 + iy_3|$. (b) The 419 boxes with bound above 0.99, with height x_3 ; the largest bounds sit in the interior of $\mathcal{R}_r$, not at its extreme points.

Upper bound. For $b \in Q$ and every h , $\text{Var}_{v_b}(\varphi_h) \leq v_b[(\varphi_h - \langle k, h \rangle)^2] \leq e^{\epsilon/8} \max_i v_{b_i}[(\varphi_h - \langle k, h \rangle)^2]$ by Lemma 33(ii).

Lower bound. Integrating the right inequality of (12) against $f \geq 0$ and dividing by $Z(b) \leq \sum_i \omega_i Z(b_i)$, the left one, gives $v_b(f) \geq e^{-\epsilon/8} \min_i v_{b_i}(f)$. With $f = (\varphi_h - v_b \varphi_h)^2$ this gives $\text{Var}_{v_b}(\varphi_h) \geq e^{-\epsilon/8} \min_i \text{Var}_{v_{b_i}}(\varphi_h)$.

Third cumulant. Put $\delta = m(b) - k$. The mean value theorem and the upper bound give $\|\delta\| \leq \tau\rho + e$, with ρ the largest distance from b_0 in Q and e the error of the computed mean. Expanding the central moment about k ,

$$D^3 g(b)(u, v, v) = T_b(u, v, v) - \delta_u \text{Var}(\varphi_v) - 2\delta_v \text{Cov}(\varphi_u, \varphi_v) - \delta_u \delta_v^2,$$

with $T_b = v_b[(\varphi - k)^{\otimes 3}]$, $\delta_u = \langle \delta, u \rangle$ and the moments under v_b . For a function G of either sign, writing $G = G^+ - G^-$ and applying the two bounds above to $G^\pm$ gives $|v_b(G)| \leq (\max_i |v_{b_i}(G)| + \epsilon \max_i v_{b_i}|G|)/(1 - \epsilon)$ with $\epsilon = 1 - e^{-\epsilon/8}$. For $G = (\varphi_u - k_u)(\varphi_v - k_v)^2$ we use $|v_{b_i}(G)| \leq \|T_{b_i}\|$, the largest singular value of the 18×324 unfolding of T_{b_i} , and $v_{b_i}|G| \leq 2\sqrt{3} v_{b_i}[(\varphi_v - k_v)^2]$, since $\|U - k\|_{\text{HS}} \leq 2\sqrt{3}$. The largest eigenvalues and singular values are bounded as in Sec. VII G, and the maxima over these boxes are the numbers stated (Appendix C). $\square$

Proof of Theorem 7. Let first $\beta_W \leq \frac{3}{4}$, so that $r = 2\beta_W \leq \frac{3}{2}$. Every field met in Lemmas 36 and 37 lies in $\frac{3}{2}K$, so the constants of Lemma 38 at $\frac{3}{2}$, $\Lambda(\frac{3}{2}) \leq \frac{0.999887}{3}$ from Proposition 34 and $\gamma = \gamma(\frac{3}{2}) = \frac{187}{128}$ apply; with these fixed constants every quantity in (15) is increasing in β_W . For $\frac{3}{4} < \beta_W \leq \frac{19}{25}$

the fields lie in $\frac{38}{25}K$, and the constants of Lemma 38 at $\frac{38}{25}$, $\Lambda(\frac{38}{25}) \leq \frac{1.029967}{3.04}$ from Proposition 35 and $\gamma(\frac{38}{25}) = \frac{4643}{3200}$ apply in the same way. Changing an active link j by η changes one staple of each of its ten partners outside its block by a matrix of norm η , so the field b_i of each such partner moves by $N\beta\eta$, and Lemma 37 gives $\sum_B C_{Bj} \leq 10N\beta K_3 = \alpha_3(\beta_W)$. In ball arithmetic, at $\beta_W = \frac{3}{4}$, $w \leq 0.017935$, $s \leq 0.34088$, $C \leq 0.043520$, $K_3 \leq 0.38609$ and $\alpha_3 \leq 0.96521$; at $\beta_W = 0.72$, $\alpha_3 \leq 0.91815$; at $\beta_W = 0.76$, with the second set of constants, $w \leq 0.018609$, $s \leq 0.34673$, $C \leq 0.045770$, $K_3 \leq 0.39454$ and $\alpha_3 \leq 0.99950$. Proposition 17, whose proof does not use the group, applies with the Hilbert–Schmidt metric. $\square$

Remark 39. Theorem 6 ends where $2r\Lambda(r)$ reaches one. The blocks remove the factor $\frac{12}{10}$ of the single-link sum and keep most of it after paying for the three links: at $\beta_W = \frac{3}{4}$ the gap bound rises from $1.1 \cdot 10^{-4}$ to 0.0177, and the block bound is the better one for $\beta_W \geq 0.715$. Before the corrections for the three links their sum is $\frac{10}{12}$ of the single-link sum $2r\Lambda(r)$, so it can stay below one while $2r\Lambda(r)$ is slightly above one; Proposition 35 certifies $2r\Lambda(r) \leq 1.03$ at $r = 1.52$, which is enough at $\beta_W = 0.76$, with little room. Evaluated in floating point with the constants of a scan of $\mathcal{R}_r$, the block sum stays below one up to $\beta_W \approx 0.78$; that computation is not certified and is not used.

IX. WHERE THE METHODS STOP

SU(2). Theorem 2 closes the single-link route in every forest gauge at $\beta_W = \frac{1}{3}$. Blocks of several links, where the

influence of a boundary link on a whole block is bounded at once, in the spirit of the Dobrushin–Shlosman criterion [31], pass it: plaquette blocks reach $\beta_W = 0.345$ (Theorem 3), and the sharper floating-point bound of Remark 18 stops near 0.357. Larger blocks are the next step.

Curvature. Remark 26 locates the end of the curvature method in this gauge between $|\beta| = \frac{1}{29}$ and $\frac{1}{26.25}$.

SU(3). The route of Sec. VII bounds the $H^{-1}(v_b)$ norm of $\varphi_h - v_b \varphi_h$, which is at least $L_h(b)$, and needs it uniformly on rK . Its Galerkin values (polynomials of degree two) at matrices of rK give $2r$ times the norm equal to 0.79 at $r = 1.25$, 0.89 at $r = 1.4$, 0.97 at $r = 1.5$ and 1.009 at $r = 1.55$, so the route stops at or below $\beta_W \approx 0.77$ even with exact constants. With $u = s\varphi_h$, Lemma 30 stops earlier: in floating point its largest values on $\mathcal{R}_r$ lie on the face $\text{diag}(r, \rho, \rho)$ with ρ between $0.75r$ and $0.8r$, and $2r$ times the bound is 0.9458 at $r = 1.36$, 0.9767 at $r = 1.40$ and 1.0080 at $r = 1.44$, so this corrector ends near $\beta_W \approx 0.715$. With $u = q_h$ of degree two the largest values on the same face are 0.9305 at $r = 1.44$, 0.9729 at $r = 1.50$ and 1.0018 at $r = 1.54$ (Fig. 7), close to the Galerkin values. Lemma 29 therefore costs little at these couplings, and most of the distance between the floating-point value at $r = 1.5$ and the certified 0.999887 of Proposition 34 is the interpolation over boxes. Between $\beta_W = \frac{3}{4}$ and 0.77 there is room only for smaller boxes. The plaquette blocks of Theorem 7 use $\Lambda(2\beta_W)$ as well, with a little more room: they reach 0.76 with the certified cover of Proposition 35 and near 0.78 in floating point (Remark 39); past that one needs larger blocks, whose sensitivity does not go through the one-link modulus.

Other groups. For groups other than SU(2) and SU(3) only the curvature route is carried out here (Theorem 5), and it stops where the bound $\frac{29}{2}$ uses up the curvature of G , at $N|\beta| = \frac{29}{2}\kappa_G$. The Dobrushin route needs a bound for the Kantorovich modulus of the one-link measure of each group, which is not attempted here.

The four open problems. Our earlier paper [14] listed four problems between the strong-coupling results and the Clay problem [47]: (1) uniqueness and a gap at intermediate fixed coupling; (2) the crossover from strong coupling to Balaban’s weak-coupling regime; (3) the continuum limit with the Osterwalder–Schrader axioms; and (4) a gap in physical units. The present paper moves the boundary of (1): for SU(2) from $\beta_W = \frac{2}{9}$ to 0.345, for SU(3) from $\frac{9}{32}$ to 0.76, with a gap at $\frac{3}{4}$ of 0.0177 in lattice units, and for every compact simple group, the exceptional groups included, past the curvature ceiling of the ungauged measure, with an explicit gap. It does not close (1); the region between these values and the scaling region, near $\beta_W \approx 2.2$ for SU(2) and $\beta_W \approx 6$ for SU(3), is open, and with it problem (2). Problems (3) and (4) are not solved here. Every bound proved here is uniform in lattice units, so by [15] none of them can by itself produce a continuum mass gap; that requires control of the gap in physical units along the renormalisation group flow [5, 17].¹³

¹³ Every one of these approaches needs an input at physical length scales. Renormalisation group approaches to the ultraviolet behaviour of gauge

The continuum problem. Two general facts limit what can be imported into problems (3) and (4) above from constructions of continuum measures. Both concern a random distribution F on $\mathbb{R}^d$, a Borel probability measure on the tempered distributions with finite moments, and the growth condition

$$\mathbb{E} e^{F(\chi)} \leq e^{C(\|\chi\|_1 + \|\chi\|_2^2)} \quad \text{for all real } \chi \in \mathcal{S}(\mathbb{R}^d), \quad (16)$$

which is the growth condition of the regularity axiom of Glimm and Jaffe at $p = 2$ [39, Sect. 6.1, (6.1.5)]. It is the hypothesis under which reflection positivity across coordinate and diagonal hyperplanes is shown in [73] to give rotation invariance of the Schwinger functions, a step proposed there for the continuum limit of lattice gauge theories.

Proposition 40. *Let F satisfy (16), and suppose that its connected two-point function is, off the diagonal, a locally integrable function $S_2(x, y)$ with $S_2(x, y) \geq c|x - y|^{-2\Delta}(\log \frac{1}{|x - y|})^{-k}$ for $0 < |x - y| \leq r_0$, where $c > 0$, $k \geq 0$ and $2\Delta > d$. Then F does not exist: no random distribution has both properties.*

Proof. Applying (16) to χ and $-\chi$ and using $\cosh x \geq 1 + \frac{x^2}{2}$ gives $\mathbb{E} F(\chi)^2 \leq 2(e^{C(\|\chi\|_1 + \|\chi\|_2^2)} - 1)$. Fix nonnegative $\varphi, \psi \in C_c^\infty(\mathbb{R}^d)$, not zero, with supports in a ball of radius $\frac{R}{2}$ and at distance $\delta > 0$ from each other, and put $\varphi_\varepsilon(x) = \varepsilon^{-d/2} \varphi(x/\varepsilon)$, and likewise ψ_ε . Then $\|\varphi_\varepsilon\|_2 = \|\varphi\|_2$ and $\|\varphi_\varepsilon\|_1 = \varepsilon^{d/2} \|\varphi\|_1 \leq \|\varphi\|_1$ for $\varepsilon \leq 1$, so the covariance of $F(\varphi_\varepsilon)$ and $F(\psi_\varepsilon)$ is bounded uniformly in ε by the Cauchy–Schwarz inequality. On the other hand the supports of φ_ε and ψ_ε are disjoint, and for u, v in the supports $\varepsilon\delta \leq |\varepsilon u - \varepsilon v| \leq \varepsilon R$. For $\varepsilon R \leq r_0$ and $\varepsilon R < 1$ the lower bound for S_2 therefore gives

$$\begin{aligned} \text{Cov}(F(\varphi_\varepsilon), F(\psi_\varepsilon)) &= \varepsilon^d \iint \varphi(u) \psi(v) S_2(\varepsilon u, \varepsilon v) du dv \\ &\geq c R^{-2\Delta} \varepsilon^{d-2\Delta} \left(\log \frac{1}{\varepsilon \delta} \right)^{-k} \|\varphi\|_1 \|\psi\|_1, \end{aligned}$$

which tends to infinity as $\varepsilon \rightarrow 0$. $\square$

The action density $\text{Tr} F_{\mu\nu} F_{\mu\nu}$ has dimension $\Delta = 4$ in four dimensions. A two-sided short-distance law of the form $|x - y|^{-8}$ times the fourth power of the running coupling, which is what asymptotic freedom¹⁴ predicts and what the conditional construction of [74] states for the smeared action density, has $2\Delta = 8 > 4$ and $k = 2$; so that field cannot satisfy (16), and rotation invariance for it has to come from another source, as

and gravitational theories, of which [5, 28] are the rigorous ones for Yang–Mills theory, have also been proposed for the effective dimension of space-time [66] and for completions of quantum gravity [11, 16]; they do not bear on the infrared estimates needed here.

¹⁴ For SU(N) the renormalisation group gives, with $b_0 = 11N/(48\pi^2)$ and the invariant scale Λ ,

$$g(a)^{-2} = 2b_0 \log \frac{1}{a\Lambda} + O(\log \log \frac{1}{a\Lambda}),$$

so the inverse coupling $\beta_W = 2N/g(a)^2$ grows like $\frac{11N^2}{12\pi^2} \log \frac{1}{a\Lambda}$ along the continuum tuning.

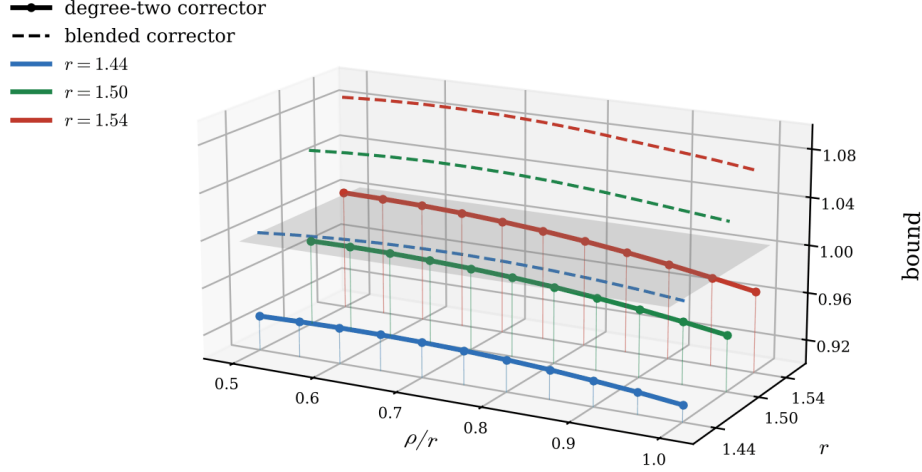


FIG. 7. The floating-point value of $2r \sup_h$ of the bound of Lemma 30 on the face $b = \text{diag}(r, \rho, \rho)$, as a function of ρ/r , for $r = 1.44, 1.50$ and 1.54 : dashed with the corrector $u = s\phi_h$ for the best s , solid with $u = q_h$ of degree two for the best C . The grey plane is the Dobrushin threshold one. The first corrector is above it already at $r = 1.44$; the second stays below it at $r = 1.50$, where Proposition 34 certifies the bound on all of $\mathcal{R}_r$, and crosses it by $r = 1.54$.

in [74], where it is reduced to a separate infrared hypothesis. In three dimensions the free part of the action density already has $2\Delta = 6 > 3$. Condition (16) suits fields of dimension at most $\frac{d}{2}$, such as the field strength of the free Maxwell field, whose covariance is bounded in momentum space.

Positivity, symmetry and regularity do not produce a gap by themselves.

Proposition 41. *On $\mathbb{R}^4$ let F be the centred Gaussian random distribution with covariance*

$$\widehat{C}(p) = \int_0^1 \frac{m^2}{|p|^2 + m^2} dm.$$

Then F is invariant under all Euclidean motions, reflection positive, satisfies (16) with $C = \frac{1}{2}$, and its Osterwalder–Schrader Hamiltonian has no gap above the vacuum.

Proof. Since $0 \leq \widehat{C} \leq 1$, $\mathbb{E}e^{F(\chi)} = e^{\frac{1}{2}\langle \chi, C\chi \rangle} \leq e^{\frac{1}{2}\|\chi\|_2^2}$, with the unitary Fourier transform. Invariance is clear. The covariance is the integral over $m \in (0, 1)$, with the weight $m^2 dm$, of the covariances of free fields of mass m , each of which is reflection positive [39, Sect. 6.2]; a positive combination of reflection positive covariances is reflection positive, and so is the Gaussian measure with such a covariance [39, Sect. 6.2]. For $\chi(t, \mathbf{x}) = h(t)g(\mathbf{x})$ with $h \geq 0$ supported in $t > 0$, not zero, and $\hat{g}(0) \neq 0$, the vector $v = F(\chi)$ is orthogonal to the vacuum, and after the Fourier transform in $\mathbf{x}$,

$$\begin{aligned} \langle v, e^{-sH} v \rangle &= \int_0^1 m^2 dm \int \frac{|\hat{g}(\mathbf{k})|^2}{2\omega} e^{-\omega s} \\ &\quad \times \left(\int_0^\infty h(t) e^{-\omega t} dt \right)^2 d\mathbf{k}, \end{aligned}$$

with $\omega = \sqrt{|\mathbf{k}|^2 + m^2}$ and constants absorbed in $\hat{g}$. If H had a gap $m_0 > 0$, the left side would be at most $e^{-m_0 s} \|v\|^2$. For

$0 < \eta < m_0$ the part of the integral with $m, |\mathbf{k}| \leq \eta/2$ is at least $c_\eta e^{-\eta s}$ with $c_\eta > 0$, a contradiction as $s \rightarrow \infty$. $\square$

The Gaussian field of Proposition 41 is not a gauge theory; it shows only that an infrared estimate, such as a bound $\langle v_a, T_a^{[t/a]} v_a \rangle \leq C_v e^{-m_* t}$ uniform in the spacing a along the continuum tuning, has to be proved in addition to the axioms. The lattice bounds of this paper are of this form only with m_* proportional to $1/a$, and by [15] such bounds give trivial limits. The routes we know to a continuum gap therefore all stop at the same two estimates: the construction of the infinite-volume continuum measure past Bałaban’s small-field analysis [5, 7], which the conditional construction of [74] assumes clause by clause and which yields there a non-negative Hamiltonian without any statement on its spectrum, and a decay estimate in physical units at large scales. The finite-size criterion described in [74, Sect. 9.4] reduces a gap to a threshold condition on one box, uniformly in the cutoff; as stated there, the condition fails in every small box and would have to be checked at large scales, which the construction does not reach.

X. THE CONTINUUM STEP AT ONE PHYSICAL SCALE

The bounds of the preceding sections hold at fixed coupling and give a gap that is at least a fixed number in lattice units, and [15] shows that no such bound can produce a continuum gap. This section replaces the single-link condition of Proposition 11 by a condition on one cube of n^4 sites, averaged over its translates in the manner of Dobrushin and Shlosman [31], and proves that the condition gives a gap of order $1/n$ in lattice units. If it holds with n proportional to $1/a$, that is, in a cube of fixed physical size, the gap survives in physical units

and passes to every continuum limit (Theorem 45). The two statements that remain unproved, the cube estimate and the existence of the limit, are stated in Sec. XI.

A. Averaged cubes

For a finite set $V \subset A$ let $\mathcal{W}(\cdot | Q)$ be the probability measure on G^V with density proportional to $\exp(N\beta \sum_{p \cap V \neq \emptyset} \text{Re Tr } Q_p^F)$ with respect to Haar measure, the links outside V being those of Q . Its conditional law of Q_e , $e \in V$, given the other links of V is γ_e^F . Write $d_V(Q_V, Q'_V) = \sum_{i \in V} \delta(Q_i, Q'_i)$ and, for $j \in A \setminus V$,

$$C_{V,j} = \sup \frac{W_1^{d_V}(\mathcal{W}(\cdot | Q), \mathcal{W}(\cdot | Q'))}{\delta(Q_j, Q'_j)},$$

the supremum over pairs that differ only at j ; $W_1^{d_V}$ is the Kantorovich distance for d_V . The coefficient vanishes unless j shares a plaquette with a link of V .

Let $n \geq 2$ be even and let V be the set of active links (x, μ) with $x \in [0, n]^4$. The direction field has period two, so the translations by $2\mathbb{Z}^4$ preserve F , A and γ^F , and the translates $V_x = V + x$, $x \in 2\mathbb{Z}^4$, satisfy $C_{V_x, j+x} = C_{V,j}$. Every active link lies in exactly $N_n = (n/2)^4$ of them. A *class* is an orbit of $2\mathbb{Z}^4$ on A , that is, a residue of the base point modulo two together with a direction; there are at most 64 classes. For a class c put

$$K_c = \sum_{j \in c \setminus V} C_{V,j}.$$

Definition 42. For $0 < \alpha < 1$, the condition $D(n, \alpha)$ at the coupling β is $K_c \leq \alpha N_n$ for every class c .

Since $\sum_{x \in 2\mathbb{Z}^4, j \notin V_x} C_{V_x, j} = \sum_x C_{V, j-x} [j-x \notin V] = K_{c(j)}$ for the class $c(j)$ of j , the condition says that the total influence of a link on the translates that do not contain it is at most the fraction α of the number of translates that do. Figure 8 shows the cubes.

For $n = 2$ it is a bound on blocks of 48 links. Distances between links are those of their midpoints; a link of V and a link sharing a plaquette with V are at distance at most $\rho_n = 2n$, and c_n denotes the number of active links within distance ρ_n of a given one.

Theorem 43. Let $\beta \geq 0$ and suppose that $D(n, \alpha)$ holds. Then statements (i)–(iv) of Proposition 11 hold, (v) holds for every $\lambda < \log(1/\alpha)/\rho_n$, and the transfer matrix of ω has spectral gap at least $\log(1/\alpha)/(2n)$.

Proof. The proof of Proposition 11 uses the coefficients $C_{ee'}$ only through Lemma 112 and Corollary 113. Lemma 115 of Appendix B gives the same conclusions under $D(n, \alpha)$, for every $\lambda > 0$ with $\alpha e^{\lambda \rho_n} < 1$, with the constant $\theta D/(1-\theta)$ replaced by $D(1 + c_n e^{\lambda \rho_n})/(1-\theta)$ and $\theta = \alpha e^{\lambda \rho_n}$. The measures used there are invariant under every $\mathcal{W}_x$ with V_x disjoint from the exceptional set: a Gibbs measure of γ^F and its conditional measures given the links of a finite set by Lemma 115(a), and the extended finite-volume measure of part (ii) because every plaquette that meets such a V_x lies in $P(\Lambda)$. Part (vi) then follows from Proposition 116 as before, with every $\lambda < \log(1/\alpha)/(2n)$. $\square$

B. The continuum limit

Let $a_k = 2^{-k} a_0$ and let $\beta_k \geq 0$ be couplings, typically with $\beta_k \rightarrow \infty$. Suppose that the lattice states ω_k at β_k exist as in Proposition 11, with transfer matrices $\mathbb{T}_k$ on the reconstructed spaces $\mathcal{H}_k$ and vacua Ω_k (Appendix F). Let $\mathcal{T}$ be a complex vector space of test functions on $\mathbb{R}^4$ supported in $\{t > 0\}$ and, for each k , let $f \mapsto O_k(f)$ be a linear map from $\mathcal{T}$ to gauge-invariant local observables supported in positive lattice times; the example to keep in mind is $O_k(f) = Z_k \sum_x a_k^4 f(a_k x) \Phi(x)$ for a local gauge-invariant field Φ , a renormalisation constant Z_k and $\Phi(x)$ the translate of Φ to x . Write ϑ for the time reflection and τ_t for the time translation by t lattice steps.

Definition 44. A *continuum limit* of (ω_k, O_k) is a Hilbert space $\mathcal{H}$ with a unit vector Ω , a self-adjoint operator $H \geq 0$ with $H\Omega = 0$, and a linear map $f \mapsto \psi(f)$ from $\mathcal{T}$ to $\mathcal{H}$ with dense range, such that for all $f, g \in \mathcal{T}$ and all dyadic $\tau \geq 0$,

$$\langle \psi(f), e^{-\tau H} \psi(g) \rangle = \lim_{k \rightarrow \infty} \omega_k(\vartheta(\overline{O_k(f)}) \tau_{\tau/a_k} O_k(g)),$$

$$\langle \Omega, \psi(f) \rangle = \lim_{k \rightarrow \infty} \omega_k(O_k(f)),$$

the limits taken over the k for which τ/a_k is an integer. It is *nontrivial* if $\mathcal{H} \neq \mathbb{C}\Omega$.

The Osterwalder–Schrader reconstruction [62, 63] produces such a limit from convergent renormalised Schwinger functions that satisfy the axioms; the definition records only what the next theorem uses.

Theorem 45. Suppose that there are $\alpha < 1$, $\ell > 0$ and even n_k with $n_k a_k \leq \ell$ such that $D(n_k, \alpha)$ holds at β_k for every k . Then every continuum limit of (ω_k, O_k) satisfies

$$\sigma(H) \subset \{0\} \cup [m_*, \infty), \quad m_* = \frac{\log(1/\alpha)}{2\ell},$$

and 0 is a simple eigenvalue. If the limit is nontrivial, $\sigma(H) \cap [m_*, \infty) \neq \emptyset$, so H has a mass gap of at least m_* .

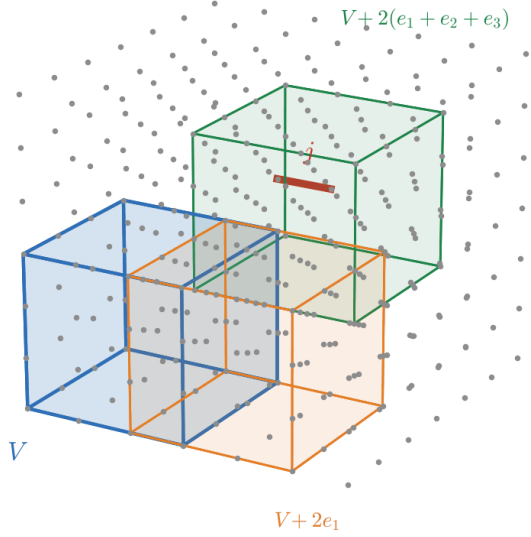
Proof. By Theorem 43 the state ω_k exists and $\mathbb{T}_k$ has gap at least $\log(1/\alpha)/(2n_k)$ on the complement of Ω_k . Let $f \in \mathcal{T}$, put $O = O_k(f)$ and $\varphi_k = \Psi_O - \omega_k(O)\Omega_k$. In the reconstruction of Appendix F, $\langle \Psi_F, \mathbb{T}_k^t \Psi_H \rangle = \omega_k(\vartheta(\bar{F}) \tau_t H)$, so $\langle \Omega_k, \Psi_O \rangle = \omega_k(O)$, and $\mathbb{T}_k \Omega_k = \Omega_k$; hence, for integers $t \geq 0$,

$$\begin{aligned} \langle \varphi_k, \mathbb{T}_k^t \varphi_k \rangle &= \omega_k(\vartheta(\bar{O}) \tau_t O) - |\omega_k(O)|^2 \\ &\leq e^{-t \log(1/\alpha)/(2n_k)} \|\varphi_k\|^2. \end{aligned}$$

Take $\tau \geq 0$ dyadic and k large, so that $t = \tau/a_k$ is an integer; then $t/(2n_k) = \tau/(2n_k a_k) \geq \tau/(2\ell)$, and the right side is at most $e^{-m_* \tau} \|\varphi_k\|^2$. Both sides converge by Definition 44. With $\varphi = \psi(f) - \langle \Omega, \psi(f) \rangle \Omega$ and $e^{-\tau H} \Omega = \Omega$ the limit reads

$$\langle \varphi, e^{-\tau H} \varphi \rangle \leq e^{-m_* \tau} \|\varphi\|^2.$$

The maps $f \mapsto O_k(f)$ and $f \mapsto \psi(f)$ are linear, so the vectors φ form a linear subspace, and it is dense in $\Omega^\perp$ because the range of ψ is dense. The inequality extends to $\Omega^\perp$ by continuity,

(a) averaged cubes, $n = 4$ 

(b) one physical scale

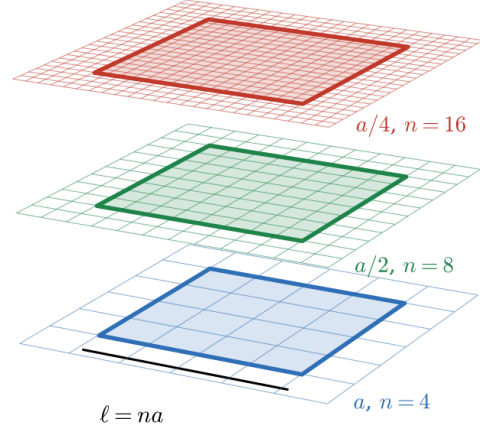


FIG. 8. (a) The cube V of side $n = 4$ and two of its translates by $2\mathbb{Z}^4$, drawn in three of the four dimensions. A link j outside V enters K_c through every translate whose boundary it touches, and $D(n, \alpha)$ compares that total with the $N_n = (n/2)^4$ translates that contain j . (b) The cube of Theorem 45: its physical side $\ell = na$ is held fixed while the spacing is halved, so n doubles at each step and the lattice gap $\log(1/\alpha)/(2n)$ is of order a/ℓ .

and from dyadic τ to all $\tau \geq 0$ by the strong continuity of $e^{-\tau H}$. For $\varphi \in \Omega^\perp$ with spectral measure μ_φ of H this says $\int e^{-\tau E} d\mu_\varphi(E) \leq e^{-m_*\tau} \mu_\varphi([0, \infty))$ for all $\tau \geq 0$; if $\mu_\varphi([0, m_* - \eta]) > 0$ for some $\eta > 0$, the left side is at least $\mu_\varphi([0, m_* - \eta])e^{-(m_* - \eta)\tau}$, which exceeds the right side for large τ . So μ_φ lives on $[m_*, \infty)$ for every $\varphi \in \Omega^\perp$, which gives both claims about 0, and the spectrum on $\Omega^\perp \neq \{0\}$ lies in $[m_*, \infty)$ and is not empty. $\square$

The proof uses $D(n_k, \alpha)$ only through the decay of the two-point functions, and the same argument accepts any decay with constants that depend on the observable.

Theorem 46. *Let $m_0 > 0$ and let $\mathcal{T}_0$ span $\mathcal{T}$. Suppose that for each $f \in \mathcal{T}_0$ there are $C_f < \infty$, $\tau_f \geq 0$ and k_f such that*

$$C_k^f(\tau) = \omega_k(\vartheta(\overline{O_k(f)}) \tau_{\tau/a_k} O_k(f)) - |\omega_k(O_k(f))|^2 \leq C_f e^{-m_0 \tau}$$

for all $k \geq k_f$ and all dyadic $\tau \geq \tau_f$ with $\tau/a_k \in \mathbb{Z}$. Then every continuum limit of (ω_k, O_k) satisfies $\sigma(H) \subset \{0\} \cup [m_0, \infty)$, and 0 is a simple eigenvalue. If the limit is nontrivial, H has a mass gap of at least m_0 .

Proof. For $f \in \mathcal{T}$ put $\varphi(f) = \psi(f) - \langle \Omega, \psi(f) \rangle \Omega$, which is linear in f . Since $e^{-\tau H}$ is self-adjoint and fixes Ω , Definition 44 gives $\langle \varphi(f), e^{-\tau H} \varphi(f) \rangle = \lim_k C_k^f(\tau)$. For $f \in \mathcal{T}_0$ let μ be the spectral measure of H at $\varphi(f)$. Then $\int e^{-\tau E} d\mu(E) \leq C_f e^{-m_0 \tau}$ for dyadic $\tau \geq \tau_f$. If $\mu([0, m_0 - \eta]) > 0$ for some $\eta \in (0, m_0)$, the left side is at least $\mu([0, m_0 - \eta])e^{-(m_0 - \eta)\tau}$, which is impossible for large τ . Hence the spectral projection $E_{[0, m_0)}$ of H vanishes on $\varphi(\mathcal{T}_0)$. By linearity it vanishes on $\varphi(\mathcal{T})$, and

since its kernel is closed, it vanishes on the closure of $\varphi(\mathcal{T})$, which is $\Omega^\perp$ because $\psi(\mathcal{T})$ is dense. So $E_{[0, m_0)} = \langle \Omega, \cdot \rangle \Omega$, which gives both claims. $\square$

Theorem 45 is the case $m_0 = m_*$, $\tau_f = 0$ and $C_f = \sup_{k \geq k_f} \|\varphi_k\|^2$, which is finite because $\|\varphi_k\|^2 = C_k^f(0)$ converges. Theorem 46 needs neither a lattice gap nor the gauge of this paper. It applies to the output of a cluster expansion of any coarse effective theory, provided the expansion bounds the correlations of a spanning set of renormalised observables in physical units.

Theorem 45 bypasses the obstruction of [15]: under the hypothesis the lattice gap is at least $\log(1/\alpha)/(2n_k)$, of order a_k/ℓ in lattice units, which tends to zero, and only its value in physical units is bounded below. In the language of [15] the hypothesis implies the gap part (R2) of Hypothesis R, with $m = m_*$, and the conclusion holds for every continuum limit without a separate assumption on the gap.

C. Smoothed observables

The observables $O_k(f)$ of Definition 44 need not be local. The renormalised fields that one expects to converge are smoothed by the gradient flow of the Wilson action [55]¹⁵; at positive flow time their correlation functions are finite to

¹⁵ The flow moves the gauge field along the negative gradient of the Wilson action. At positive flow time the flowed fields are smooth on the scale of the square root of the flow time, and in perturbation theory their gauge-invariant composite fields need no further renormalisation [55, 56].

all orders of perturbation theory once the coupling is renormalised [56]. Flow in all four directions mixes the two sides of the reflection plane, but the spatial flow of [67, Sec. II.B] does not: it flows the links of each time slice under the Wilson action of that slice and leaves the temporal links alone.

Fix a bi-invariant Riemannian metric on G . For a configuration U of all links of $\mathbb{Z}^4$ (for the states of this paper, the links of F equal to 1), let $\Phi_s(U)$, $s \in \mathbb{R}$, solve

$$\frac{d}{ds} U_e = -\nabla_e \sum_{p \ni e, p \text{ spatial}} (N - \text{Re Tr } U_p)$$

for every spatial link e , the sum running over the four plaquettes of the time slice of e that contain it, with the temporal links held fixed and ∇_e the gradient in the variable U_e .

Proposition 47. (i) Φ_s is defined on every configuration of the infinite lattice, $\Phi_s \circ \Phi_{s'} = \Phi_{s+s'}$, and Φ_s commutes with gauge transformations, with spatial and temporal translations and with the reflections in time. The link $\Phi_s(U)_e$ depends only on the spatial links of the time slice of e , and there are constants κ, c depending only on G such that, if Φ_s^R denotes the flow of the links of that slice within distance R of e with all other links held fixed,

$$\delta(\Phi_s(U)_e, \Phi_s^R(U)_e) \leq D e^{\kappa|s|} \frac{(c|s|)^{R+1}}{(R+1)!}$$

for every configuration U .

(ii) Let ω be a state as in Appendix F and O a bounded gauge-invariant function of finitely many links at times in $[t_0, t_1]$, continuous in them. Then $O \circ \Phi_s$ is a bounded gauge-invariant function of the links at times in $[t_0, t_1]$ and a uniform limit of gauge-invariant local functions at those times.

(iii) Theorem 45 holds with $O_k(f)$ replaced by $O_k(f) \circ \Phi_{s_k}$ for any $s_k \in \mathbb{R}$, in particular for $s_k = \tau/a_k^2$ at a fixed flow time τ in physical units.

(iv) The image measure $(\Phi_s)_* \omega$ is reflection positive, and its transfer matrix is unitarily equivalent to that of ω , the unitary fixing the vacuum.

Proof. (i) Embed G in a matrix algebra. The right side of the equation is a smooth function of the links of the four plaquettes containing e , tangent to G at U_e , so on the Banach space of bounded assignments of matrices to the spatial links it is a vector field that is bounded and Lipschitz on a neighbourhood of the configurations with values in G , with a Lipschitz constant that does not depend on the volume because every link meets finitely many others. The Picard–Lindelöf theorem gives a unique local solution; it stays on G because the field is tangent there, and since the field is bounded on G the solution exists for all s ; uniqueness gives the group law. The action of a time slice and the metric are invariant under gauge transformations, translations and the reflections, so the flow commutes with them by uniqueness. Since the equation of a spatial link involves only links of its own slice, so does the solution. For the comparison, let $d_{e'}(s)$ be the distance between the two flows at a link e' of the slice. Smoothness on the compact group gives constants κ_0, c_1 with $d_{e'}(s) \leq \int_0^{|s|} (\kappa_0 d_{e'} + c_1 \sum_{e'' \sim e'} d_{e''})$ for e'

within distance R of e , where $e'' \sim e'$ means that the two links share a spatial plaquette (twelve neighbours), while $d_{e'} \leq D$ beyond. Iterating the integral form of this inequality, a term that reaches the links beyond distance R needs at least $R+1$ steps of $\sim$, and $\sum_{k>R} x^k/k! \leq e^x x^{R+1}/(R+1)!$; this gives the bound with $c = 12c_1$ and $\kappa = \kappa_0 + c$.

(ii) The truncated flow Φ_s^R commutes with gauge transformations for the same reason as Φ_s , so $O \circ \Phi_s^R$ is a gauge-invariant function of finitely many links at times in $[t_0, t_1]$. By (i) and the uniform continuity of O , it converges to $O \circ \Phi_s$ uniformly as $R \rightarrow \infty$.

(iii) The proof of Theorem 45 uses $O_k(f)$ only through the vectors $\Psi_{O_k(f)}$ and the identity $\langle \Psi_F, \mathbb{T}^t \Psi_H \rangle = \omega(\vartheta(\bar{F}) \tau_t H)$. By (ii) and the footnote in Appendix F, both extend to uniform limits of local functions at positive times, and $f \mapsto O_k(f) \circ \Phi_{s_k}$ is linear.

(iv) By (i), Φ_s and Φ_{-s} preserve the functions of the links at positive times and commute with ϑ and τ_t . Hence $(\Phi_s)_* \omega(\vartheta(\bar{F}) \tau_t H) = \omega(\vartheta(\bar{F} \circ \Phi_s) \tau_t (H \circ \Phi_s))$, which for $t = 0$ and $H = F$ is nonnegative, and $F \mapsto F \circ \Phi_s$ induces an isometry between the reconstructed spaces that intertwines the transfer matrices and fixes the vacuum; Φ_{-s} gives its inverse. $\square$

Part (iv) says that smoothing changes the variables and not the theory: it neither creates nor removes a gap. Its use is in the choice of the observables. With $O_k(f)$ the spatially flowed fields at a fixed physical flow time, for instance $Z_k \sum_x a_k^4 f(a_k x) (\mathcal{E}_\tau(x) - \omega_k(\mathcal{E}_\tau(x)))$ with $\mathcal{E}_\tau(x)$ the flowed action density of the slice, statement (E) asks for the convergence of correlation functions that are known to be finite in perturbation theory [56], and Theorem 45 applies to them. Theorem 48 below constructs their continuum limits along subsequences; their nontriviality is not proved here.

D. Renormalised flowed observables

The spatially flowed fields can be renormalised by one condition at a single physical separation, and their correlation functions then have continuum limits along subsequences without further assumptions. What remains open is whether such a limit is trivial, and that question reduces to one ratio of lattice two-point functions.

Fix a bounded gauge-invariant function Φ of finitely many spatial links of the slice $\{x_0 = 0\}$, continuous in them, a flow time $\tau > 0$ in physical units, and a nonzero test function $f_0 \in C_c(\mathbb{R}^4)$ supported in $\{t > 0\}$. Put $s_k = \tau/a_k^2$, write Φ_x for the translate of Φ to $x \in \mathbb{Z}^4$, and let

$$A_k = \sum_{x \in \mathbb{Z}^4} a_k^4 f_0(a_k x) (\Phi_x \circ \Phi_{s_k} - \omega_k(\Phi_x \circ \Phi_{s_k})),$$

$$C_k(\sigma) = \omega_k(\vartheta(\bar{A}_k) \tau_{\sigma/a_k} A_k)$$

for dyadic $\sigma \geq 0$ (the sum is finite, and σ/a_k is an integer for large k). By reflection positivity $C_k(0) \geq 0$; assume $C_k(0) > 0$ and set $Z_k = C_k(0)^{-1/2}$. Let $\mathcal{D}$ be the set of points $y = (t, y') \in \mathbb{R}^4$ with dyadic coordinates (in units of a_0) and $t > 0$, and for

$y \in \mathcal{D}$ and k with $y/a_k \in \mathbb{Z}^4$ let A_k^y be the translate of A_k by y/a_k .

Theorem 48. (i) For $y, y' \in \mathcal{D}$ and every k with $y/a_k, y'/a_k \in \mathbb{Z}^4$,

$$Z_k^2 |\omega_k(\vartheta(\bar{A}_k^y)A_k^{y'})| \leq 1.$$

(ii) Every sequence of k has a subsequence along which $K(y, y') = \lim Z_k^2 \omega_k(\vartheta(\bar{A}_k^y)A_k^{y'})$ exists for all $y, y' \in \mathcal{D}$ and $\rho(\sigma) = \lim C_k(\sigma)/C_k(0)$ exists for all dyadic $\sigma > 0$. Along it there is a probability measure μ on $[0, \infty]$ with

$$\rho(\sigma) = \int_{[0, \infty]} e^{-\sigma E} d\mu(E) \quad (\sigma > 0),$$

the integrand being 0 at $E = \infty$.

(iii) Along such a subsequence, let $\mathcal{T}$ be spanned by the translates $f_0(\cdot - y)$, $y \in \mathcal{D}$, and one more function $f_* \in C_c(\mathbb{R}^4)$ supported in $\{t > 0\}$ outside their span, and let O_k be the linear map with $O_k(f_0(\cdot - y)) = Z_k A_k^y$ and $O_k(f_*) = 1$. Then (ω_k, O_k) has a continuum limit in the sense of Definition 44, on which the dyadic spatial translations act by unitary operators commuting with H .

(iv) This limit is nontrivial if and only if $\mu([0, \infty)) > 0$, if and only if $\rho(\sigma) > 0$ for one dyadic $\sigma > 0$, and then $\rho(\sigma) > 0$ for all of them. Consequently a nontrivial limit exists along some subsequence if and only if $\limsup_k C_k(\sigma)/C_k(0) > 0$ for some dyadic $\sigma > 0$.

(v) If the hypothesis of Theorem 45 holds, then μ is carried by $[m_*, \infty]$ and $\rho(\sigma) \leq e^{-m_* \sigma}$.

Proof. By Proposition 47(ii) each $\Phi_x \circ \Phi_{s_k}$, and hence each A_k^y , is a uniform limit of gauge-invariant local functions at times ≥ 1 , so the reconstruction of Appendix F applies to them (its footnote). Write $\Psi_k = Z_k \Psi_{A_k}$. Since $\langle \Omega_k, \Psi_{A_k} \rangle = \omega_k(A_k) = 0$ and $\|\Psi_k\|^2 = Z_k^2 C_k(0) = 1$, Ψ_k is a unit vector orthogonal to Ω_k . The relation $\langle \Psi_F, \mathbb{T}_k^t \Psi_H \rangle = \omega_k(\vartheta(\bar{F})\tau_t H)$ for all F gives $\Psi_{\tau_t H} = \mathbb{T}_k^t \Psi_H$. Spatial translations commute with ϑ and with τ_t and leave ω_k invariant, so they act on the reconstructed space by unitaries $U_k(x)$ that commute with $\mathbb{T}_k$. Hence, for $y = (t, y')$,

$$Z_k \Psi_{A_k^y} = U_k(y'/a_k) \mathbb{T}_k^{t/a_k} \Psi_k.$$

(i) The left side of the inequality is $|\langle Z_k \Psi_{A_k^y}, Z_k \Psi_{A_k^{y'}} \rangle| \leq 1$,

because $\|U_k\| = 1$ and $0 \leq \mathbb{T}_k \leq 1$.

(ii) The set $\mathcal{D} \times \mathcal{D}$ is countable and the numbers are bounded by (i), so a diagonal subsequence gives all the limits K ; the limits $\rho(\sigma)$ are among them, since $C_k(\sigma)/C_k(0) = \langle \Psi_k, \mathbb{T}_k^{\sigma/a_k} \Psi_k \rangle$ is the value of K at $y = y' = (\sigma/2, 0)$. Let ν_k be the spectral measure (Lemma 117) of $\mathbb{T}_k$ at Ψ_k , a probability measure on $[0, 1]$, and μ_k its image under $\lambda \mapsto -a_k^{-1} \log \lambda \in [0, \infty]$. Then $C_k(\sigma)/C_k(0) = \int e^{-\sigma E} d\mu_k(E)$. The space $[0, \infty]$ is compact, so a further subsequence has $\mu_k \rightarrow \mu$ weakly (Lemma 118), and $E \mapsto e^{-\sigma E}$ is continuous on $[0, \infty]$ for $\sigma > 0$.

(iii) A finite combination $\sum_j c_j f_0(\cdot - y_j)$ with distinct y_j that vanishes has Fourier transform $\hat{f}_0(p) \sum_j c_j e^{-ip \cdot y_j} = 0$; since $\hat{f}_0$ is analytic and not zero, the exponential sum vanishes on an

open set and every c_j is zero. So the translates are linearly independent and O_k is well defined. On the space V of finite sums $v = \sum_y c_y \delta_y$ over $\mathcal{D}$ put $B(v, w) = \sum_y \bar{c}_y c'_y K(y, y')$. It is the limit of $\langle v_k, w_k \rangle$ with $v_k = \sum_y c_y Z_k \Psi_{A_k^y}$, hence positive semidefinite; let $\mathcal{H}_0$ be the completion of V modulo its null space. For dyadic $\sigma \geq 0$ and dyadic spatial x define $T(\sigma) \delta_y = \delta_{y+(\sigma, 0)}$ and $U(x) \delta_y = \delta_{y+(0, x)}$. At level k these are $\mathbb{T}_k^{\sigma/a_k}$ and $U_k(x/a_k)$, so in the limit $B(T(\sigma)v, w) = B(v, T(\sigma)w)$, $B(T(\sigma)v, T(\sigma)v) \leq B(v, v)$, $B(v, T(\sigma)v) \geq 0$, and $U(x)$ preserves B . Thus $T(\sigma)$ extends to commuting positive contractions on $\mathcal{H}_0$ with $T(\sigma)T(\sigma') = T(\sigma + \sigma')$, and $U(x)$ to unitaries commuting with them. For dyadic $\sigma > 0$ the kernel of $T(\sigma) = T(\sigma/2)^2$ equals that of $T(\sigma/2)$, so all these kernels coincide; and every δ_y with $y = (t, y')$ equals $T(t/2) \delta_{(t/2, y')}$, a vector in the range of a positive operator and so orthogonal to its kernel. Hence each $T(\sigma)$ is injective on $\mathcal{H}_0$, and $H_0 = -\log T(1)$ is a self-adjoint operator with $H_0 \geq 0$ and $e^{-\sigma H_0} = T(\sigma)$ for dyadic σ , by the uniqueness of positive roots. Put $\mathcal{H} = \mathbb{C}\Omega \oplus \mathcal{H}_0$, $H = 0 \oplus H_0$, $\psi(f_0(\cdot - y)) = \delta_y$ and $\psi(f_*) = \Omega$. The range of ψ is dense, and the identities of Definition 44 hold: for two translates they are the limits defining K , with $\tau_{\sigma/a_k} A_k^{y'} = A_k^{y'+(\sigma, 0)}$; and for f_* they read $\omega_k(1) = 1$ and $\omega_k(\vartheta(\bar{A}_k^y)) = \omega_k(\tau_{\sigma/a_k} A_k^y) = 0$.

(iv) The limit is nontrivial exactly when $\mathcal{H}_0 \neq 0$, that is when $\|\delta_y\|^2 = K(y, y) = \rho(2t)$ is positive for some $y = (t, y')$. By (ii) and Lemma 119(b), $\rho(\sigma) > 0$ for one $\sigma > 0$ if and only if $\mu([0, \infty)) > 0$, and then $\rho(\sigma) > 0$ for every $\sigma > 0$. For the last statement choose a subsequence along which $C_k(\sigma)/C_k(0)$ tends to its limsup, and refine it by (ii).

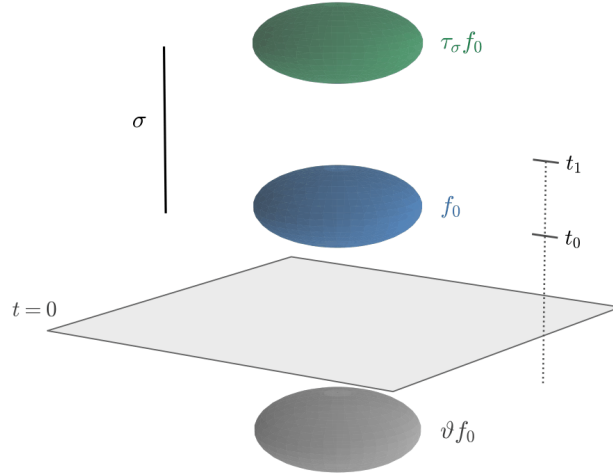
(v) Under that hypothesis the proof of Theorem 45 gives $\nu_k([0, e^{-\log(1/\alpha)/(2n_k)}]) = 1$ because $\Psi_k \perp \Omega_k$, so μ_k is carried by $[\log(1/\alpha)/(2n_k a_k), \infty] \subset [m_*, \infty]$, and so is the weak limit μ . $\square$

The measure μ is the spectral measure of the renormalised flowed field in the limit: its part on $[0, \infty)$ is what survives as a state of the continuum theory, and its atom at ∞ is the weight that the field puts on energies that diverge in physical units. The theorem constructs the renormalised continuum observables and their reconstruction for every tuning, and it reduces their nontriviality to the lattice inequality

$$\limsup_{k \rightarrow \infty} \frac{C_k(\sigma)}{C_k(0)} > 0$$

for one physical separation σ : the flowed two-point function of the smeared packet at a fixed physical distance must not become negligible against its reflected norm $C_k(0)$. If the lattice gap stayed above some $m > 0$ in lattice units, the ratio would be at most $e^{-m\sigma/a_k}$ and the limit would be trivial, so the inequality is a condition on the tuning $\beta_k \rightarrow \infty$; it is not proved here. Nor are the independence of the limit from the subsequence, Euclidean invariance, or bounds on products of more than two renormalised fields, for which (i) gives nothing because Z_k may diverge (Theorem 56 reduces them to tree bounds); (E) requires all of these.

The denominator $C_k(0)$ is the reflected norm of the packet smeared by f_0 , not a coincidence limit of point fields, and the



$$C_k(\sigma) = \langle \Psi_{A_k}, \mathbb{T}_k^{\sigma/a_k} \Psi_{A_k} \rangle: \text{the reflected packet } \vartheta f_0 \text{ paired with the packet translated by } \sigma$$

FIG. 9. The packet f_0 of Theorem 48, supported in the time window $[t_0, t_1]$ above the reflection plane $t = 0$, its reflection ϑf_0 and its translate $\tau_\sigma f_0$ by σ in time. The correlation $C_k(\sigma)$ pairs the reflected packet with the translated one; $C_k(0)$ is the reflected norm that renormalises the packet.

distinction matters. For the free massless scalar field in four dimensions with spatial momenta cut off at Λ , the covariance at spatial separation 0 is $(1 - (1 + \Lambda t)e^{-\Lambda t})/(4\pi^2 t^2)$ at time $t > 0$ and $\Lambda^2/(8\pi^2)$ at coincidence, so its ratio $2(1 - (1 + \Lambda t)e^{-\Lambda t})/(\Lambda t)^2$ tends to zero as $\Lambda \rightarrow \infty$, although the limiting field is the nontrivial Gaussian field; for packets smeared by f_0 the reflected norms of that field have finite positive limits and the ratio stays positive. The criterion of Theorem 48(iv) is about the smeared packets A_k only.

When the ratio is bounded below, renormalising at the separation σ instead of 0 removes the escape to infinite energy altogether, and one slab of the transfer matrix gives a sufficient condition for the bound. Write μ_k for the measures in the proof of Theorem 48.

Proposition 49. (i) Suppose $C_k(\sigma) \geq c C_k(0)$ for a dyadic $\sigma > 0$, a $c > 0$ and every k of a sequence. Let μ'_k be the spectral measure, in energies $E = -a_k^{-1} \log \lambda$, of the unit vector $C_k(\sigma)^{-1/2} \Psi_{A'_k}$ with $A'_k = \tau_{\sigma/2a_k} A_k$. Then μ'_k is a probability measure on $[0, \infty)$,

$$\mu'_k([\Lambda, \infty)) \leq c^{-1} e^{-\sigma \Lambda}, \quad \int E^m d\mu'_k \leq c^{-1} \left(\frac{m}{e\sigma} \right)^m,$$

and $t \mapsto C_k(\sigma + t)/C_k(\sigma)$ is Lipschitz on $[0, \infty)$ with constant $(e\sigma c)^{-1}$. A subsequence of the μ'_k converges weakly to a probability measure on $[0, \infty)$ with the same bounds.

(ii) Suppose $\langle \Psi_k, (1 - \mathbb{T}_k) \Psi_k \rangle \leq C a_k$ for a constant C and every k of a sequence. Then $C_k(\sigma)/C_k(0) \geq e^{-2C\sigma}$ whenever $C a_k \leq \frac{1}{2}$ and σ/a_k is an integer, and $\mu_k([\Lambda, \infty)) \leq C(a_k + \Lambda^{-1})$ for every $\Lambda > 0$.

Proof. (i) We have $\Psi_{A'_k} = \mathbb{T}_k^{\sigma/2a_k} \Psi_{A_k}$, so the measure μ'_k is $e^{-\sigma E} d\mu_k(E)$ divided by its mass $C_k(\sigma)/C_k(0) \geq c$, and $e^{-\sigma E}$ vanishes at $E = \infty$. For $E \geq \Lambda$ we have $e^{-\sigma E} \leq e^{-\sigma \Lambda}$, and $E^m e^{-\sigma E} \leq (m/e\sigma)^m$ for all $E \geq 0$. The ratio in question is the integral of e^{-tE} against μ'_k , and the inequality $|e^{-tE} - e^{-sE}| \leq |t - s|E$ with the moment bound for $m = 1$ gives the Lipschitz bound. The tail bound makes the sequence tight on $[0, \infty)$, so Prokhorov's theorem gives the weak limit, and the bounds pass to it because $\{E > \Lambda\}$ is open and E^m is a monotone limit of bounded continuous functions.

(ii) Let ν_k be the spectral measure of $\mathbb{T}_k$ at Ψ_k on $[0, 1]$, a probability measure, and $n = \sigma/a_k$. By Jensen's inequality for the convex function λ^n ,

$$\frac{C_k(\sigma)}{C_k(0)} = \int \lambda^n d\nu_k \geq \left(\int \lambda d\nu_k \right)^n \geq (1 - C a_k)^n \geq e^{-2C a_k n},$$

since $\log(1 - x) \geq -2x$ on $[0, \frac{1}{2}]$. Moreover $C a_k \geq \int (1 - e^{-a_k E}) d\mu_k \geq (1 - e^{-a_k \Lambda}) \mu_k([\Lambda, \infty))$, and $1 - e^{-x} \geq x/(1 + x)$. $\square$

Part (ii) asks for an upper bound on a one-slab Dirichlet form, the opposite direction to a mass gap, which bounds such forms from below. In the box of Proposition 50, with g smooth, the time smearing supplies the factor a_k up to one ratio: in the notation of its proof, $\mathbb{T}_k = A$ and $\Psi_{A_k} = \sum_n w_n \mathbb{T}_k^n u$ with $w_n = a_k g(a_k n)$ and $u = a_k^3 F^c \psi_0$, and the identity $(1 - \mathbb{T}_k) \Psi_{A_k} = \sum_n (w_n - w_{n-1}) \mathbb{T}_k^n u$ gives $\langle \Psi_k, (1 - \mathbb{T}_k) \Psi_k \rangle \leq a_k \|g'\|_{L^1} \|u\| / \|\Psi_{A_k}\|$, so the hypothesis of (ii) holds

when the vector u of a single slice stays comparable to the time-smeared packet. That comparison is not proved; Proposition 51 below bounds the ratio without it.

A one-slab bound of this kind also limits the mass that $D(n_k, \alpha)$ can supply. Let e^{-m_k} be the norm of $\mathbb{T}_k$ on $\Omega_k^\perp$. If a unit vector $\varphi_k \perp \Omega_k$ has $\langle \varphi_k, (1 - \mathbb{T}_k)\varphi_k \rangle \leq \kappa a_k < 1$, then the spectral measure of $\mathbb{T}_k$ at φ_k lives on $[0, e^{-m_k}]$, so $1 - e^{-m_k} \leq \kappa a_k$ and $m_k \leq -\log(1 - \kappa a_k)$. If in addition $D(n_k, \alpha)$ holds with $n_k a_k \leq \ell$, Theorem 43 gives $m_k \geq \log(1/\alpha)/(2n_k) \geq a_k \log(1/\alpha)/(2\ell)$, and therefore

$$\frac{\log(1/\alpha)}{2\ell} \leq \frac{-\log(1 - \kappa a_k)}{a_k} = \kappa + O(a_k).$$

So the physical mass m_* of Theorem 45 is at most the constant κ of any one-slab bound. The two hypotheses are compatible because both are statements in physical units. The one-slab bound forces $m_k \rightarrow 0$, so it cannot come from an expansion whose convergence gives a lattice gap bounded below in lattice units.

In a periodic spatial box the assumption $C_k(0) > 0$ is a theorem. Let $G = \text{SU}(N)$ with its defining representation, and replace $\mathbb{Z}^4$ by $\mathbb{Z} \times (\mathbb{Z}/L_k\mathbb{Z})^3$ with $L_k \geq 3$ and ω_k by the infinite-time Wilson state of the box. The proof of Theorem 48 uses only reflection positivity and the transfer matrix once the spatial translations are dropped, and parts (i)–(iv) hold in the box with $\mathcal{D}$ replaced by its points $(t, 0)$, the time translates of f_0 being independent by the same Fourier argument in t .

Proposition 50. *In the box, let $f_0(t, x) = g(t)h(x)$ with $g, h \geq 0$, and let $\Phi = \sum_{i < j} (N - \text{Re Tr } Q_{ij})$ be the energy of the three spatial plaquettes at the origin. If $g(a_k n)h(a_k x) > 0$ for some $n \geq 1$ and some x , then $C_k(0) > 0$.*

Proof. Lüscher proved that the transfer matrix is strictly positive [54]; we recall the argument because the proof uses it in this form.

The one-link kernel. Write χ for the character of the defining representation and $b = \beta_W > 0$. Then

$$e^{\frac{b}{N} \text{Re } \chi} = \sum_{m \geq 0} \frac{(b/2N)^m}{m!} (\chi + \bar{\chi})^m,$$

and $(\chi + \bar{\chi})^m$ is the character of $(\rho \oplus \bar{\rho})^{\otimes m}$. Every irreducible representation σ occurs in one of these tensor powers: their matrix coefficients span the algebra of polynomials in the entries of q and $\bar{q}$, which contains the constants and separates points, so it is dense in $C(G)$ by the Stone–Weierstrass theorem, while the matrix coefficients of a σ that never occurred would be orthogonal to it by Schur orthogonality. So the coefficient of every irreducible character in $e^{\frac{b}{N} \text{Re } \chi}$ is positive, and convolution with this function acts on each isotypic part of $L^2(G)$ by a positive scalar.

The transfer matrix. In the temporal gauge the transfer matrix of the box acts on the gauge-invariant part of $L^2(G^E)$, E the spatial links of one slice, as MKM , where K is the tensor product over E of the convolutions above and M is multiplication by the exponential of half the spatial plaquette weight [54, 64]. Both commute with gauge transformations,

K is positive and injective by the first step, and M is positive and invertible, so $\langle v, MKMv \rangle = \langle Mv, KMv \rangle > 0$ for $v \neq 0$. Its kernel is a continuous positive function, and K is compact, so its largest eigenvalue λ_0 is simple with an eigenfunction $\psi_0 > 0$, and with $A = MKM/\lambda_0$, a function F of the slice at time $n \geq 0$ has $\Psi_F = A^n F \psi_0$ in the reconstruction, Ω being ψ_0 .

Conclusion. Let $F = \sum_x h(a_k x) \Phi_x \circ \Phi_{s_k}$ and $F^c = F - \omega_k(F)$. Then $\Psi_{A_k} = a_k^4 p(A) F^c \psi_0$ with $p(z) = \sum_{n \geq 1} g(a_k n) z^n$, a polynomial with nonnegative coefficients, not all zero, so $p > 0$ on $(0, 1]$ and $p(A)$ is injective. The function $\sum_x h(a_k x) \Phi_x$ vanishes at the configuration with every link 1 and is positive at one with a single nontrivial link at a site where $h(a_k x) > 0$; the flow Φ_{s_k} is a bijection, so F is not constant, and as F is continuous and $\psi_0 > 0$, $F^c \psi_0 \neq 0$. Hence $C_k(0) = \|\Psi_{A_k}\|^2 > 0$. $\square$

The proposition removes the one assumption of Theorem 48 in a box; it says nothing about the ratio of part (iv), on which nontriviality depends.

The ratio of Theorem 48(iv) can be compared with correlations of a single slice field. Suppose that $f_0(t, x) = g(t)h(x)$ with $g \geq 0$ supported in $[t_0, t_1]$, $t_0 < t_1$ positive and dyadic, let

$$U_k = \sum_{x \in \mathbb{Z}^3} a_k^3 h(a_k x) (\Phi_{(0,x)} \circ \Phi_{s_k} - \omega_k(\Phi_{(0,x)} \circ \Phi_{s_k}))$$

be the centred field of the slice at time 0, smeared in space only, and put $u_k = \Psi_{U_k}$ and

$$D_k(r) = \langle u_k, \mathbb{T}_k^{r/a_k} u_k \rangle = \omega_k(\vartheta(\bar{U}_k) \tau_{r/a_k} U_k)$$

for $r \in a_k \mathbb{Z}$, $r \geq 0$. Let $n_0 = \lceil t_0/a_k \rceil$ and $n_1 = \lfloor t_1/a_k \rfloor$.

Proposition 51. *Assume $C_k(0) > 0$, and let $\sigma \geq 0$ with $\sigma/a_k \in \mathbb{Z}$.*

(i) $D_k(r) > 0$ for every r , and

$$\frac{D_k(2a_k n_1 + \sigma)}{D_k(2a_k n_0)} \leq \frac{C_k(\sigma)}{C_k(0)} \leq \frac{D_k(2a_k n_0 + \sigma)}{D_k(2a_k n_1)}.$$

(ii) $\log D_k$ and $\log C_k$ are convex and nonincreasing on $a_k \mathbb{Z}_{\geq 0}$. If $D_k(r') \geq c D_k(r)$ with $r < r'$ and $c \leq 1$, then $D_k(\beta) \geq c^{\max(1, (\beta - \alpha)/(r' - r))} D_k(\alpha)$ whenever $r \leq \alpha < \beta$; if $C_k(\sigma) \geq c C_k(0)$ with $\sigma > 0$, then $C_k(\sigma') \geq c^{\sigma'/\sigma} C_k(0)$ for $\sigma' \geq \sigma$.

(iii) Let k run through a sequence. If $D_k(r') \geq c D_k(r)$ for dyadic $0 < r < r'$ with $r \leq 2t_0$, $a \in (0, 1]$ and every large k , then for every dyadic $\sigma \geq 0$ and every large k

$$C_k(\sigma) \geq c^{\max(1, (2t_1 - 2t_0 + \sigma)/(r' - r))} C_k(0).$$

Conversely, if $C_k(\sigma) \geq c C_k(0)$ for a dyadic $\sigma > 0$ and every k , then $D_k(2t_0 + \sigma') \geq c^{\sigma'/\sigma} D_k(2t_1)$ for every dyadic $\sigma' \geq \sigma$ with $\sigma' > 2(t_1 - t_0)$ and every large k .

Proof. For a function F of the slice at time 0 and an integer $n \geq 0$ we have $\Psi_{\tau_n F} = \mathbb{T}_k^n \Psi_F$, since both have inner product $\omega_k(\vartheta(\bar{G}) \tau_n F)$ with every Ψ_G . By time translation invariance of ω_k , $A_k = \sum_n w_n \tau_n U_k$ with $w_n = a_k g(a_k n) \geq 0$, and $w_n = 0$

unless $n_0 \leq n \leq n_1$; so $\Psi_{A_k} = \sum_n w_n \mathbb{T}_k^n u_k$. Let ν be the spectral measure of $\mathbb{T}_k$ at u_k , on $[0, 1]$, and $W(\lambda) = \sum_n w_n \lambda^n$. Then

$$C_k(\sigma) = \int W(\lambda)^2 \lambda^{\sigma/a_k} d\nu(\lambda), \quad D_k(r) = \int \lambda^{r/a_k} d\nu(\lambda).$$

(i) Since $t_0 > 0$, $n_0 \geq 1$ and $W(0) = 0$; so $C_k(0) > 0$ forces $\nu((0, 1]) > 0$, and then $D_k(r) > 0$. For $0 \leq \lambda \leq 1$, $\|w\|_1 \lambda^{n_1} \leq W(\lambda) \leq \|w\|_1 \lambda^{n_0}$; inserting these bounds in $C_k(\sigma)$ and $C_k(0)$ gives the two inequalities.

(ii) Convexity of $j \mapsto \log \int \lambda^j d\rho$ for a positive measure ρ on $[0, 1]$ is Hölder's inequality, and the integrals decrease in j because $\lambda \leq 1$; take $\rho = \nu$ and $\rho = W^2 \nu$. If $\beta \leq r'$, monotonicity gives $D_k(\beta) \geq D_k(r') \geq c D_k(r) \geq c D_k(\alpha)$. If $\beta > r'$, the difference quotients of the convex function $\log D_k$ increase in both end points, so $(\log D_k(\beta) - \log D_k(\alpha))/(\beta - \alpha) \geq (\log D_k(r') - \log D_k(r))/(r' - r) \geq \log c/(r' - r)$. The statement on C_k is the same comparison of quotients on $[0, \sigma]$ and $[0, \sigma']$.

(iii) For large k the times $2t_0, 2t_1$ are multiples of a_k , and $2a_k n_0 \geq 2t_0, 2a_k n_1 \leq 2t_1$. Since D_k is nonincreasing, (i) gives

$$\frac{D_k(2t_1 + \sigma)}{D_k(2t_0)} \leq \frac{C_k(\sigma)}{C_k(0)} \leq \frac{D_k(2t_0 + \sigma)}{D_k(2t_1)}.$$

The first claim is the left inequality with (ii) for $\alpha = 2t_0 \geq r$ and $\beta = 2t_1 + \sigma$. For the second, (ii) gives $C_k(\sigma') \geq c^{\sigma'/\sigma} C_k(0)$, and the right inequality at σ' gives the claim, the two times being ordered because $\sigma' > 2(t_1 - t_0)$. $\square$

So for the packet $g \otimes h$ the open input of Theorem 48(iv) follows from an inequality between two correlations of one slice field,

$$\liminf_k D_k(r')/D_k(r) > 0 \quad \text{for one pair } 0 < r < r' \text{ with } r \leq 2t_0,$$

and by (iii) a uniform bound for one packet gives one for every packet $g' \otimes h$ with g' supported in $[t_1, \infty)$. The condition $r \leq 2t_0$ cannot be dropped: a bound at a later pair says nothing about an earlier packet, as $D_k(r) = (1 - 4^{-k})2^{-kr} + 4^{-k}2^{-r}$ shows, with $D_k(4)/D_k(3) \rightarrow \frac{1}{2}$ and $D_k(2)/D_k(1) \rightarrow 0$. A bound at a later pair does apply to the packet shifted forward in time, renormalised by its own reflected norm. The left inequality of (i) bounds the ratio without the comparison of u with the packet that was needed after Proposition 49(ii). Every bound on such ratios proved in this paper, at strong coupling or through Theorem 45, is an upper bound of the form $e^{-m(r'-r)/a_k}$; the inequality above is the opposite one and says that the slice field does not put its weight on energies that diverge in physical units. Reflection positivity and positivity of the transfer matrix do not give it, since a lattice theory with a fixed mass in lattice units has both and its ratio tends to zero. It holds for the free field.

Remark 52. Let D_k be defined as above for the massless free lattice field on $a_k \mathbb{Z}^4$ in place of Φ , with $h \in C_c^\infty(\mathbb{R}^3)$ and the lattice heat flow for a physical time $t_f > 0$ in place of the gradient flow. With $\hat{p}^2 = \sum_i (2a_k^{-1} \sin(a_k p_i/2))^2$, $\cosh(a_k \omega) = 1 + a_k^2 \hat{p}^2/2$ and $\hat{h}_k(p) = \sum_x a_k^3 h(a_k x) e^{-ia_k p \cdot x}$,

$$D_k(r) = \int_{[-\pi/a_k, \pi/a_k]^3} |\hat{h}_k(p)|^2 e^{-2t_f \hat{p}^2} \frac{e^{-r\omega}}{2a_k^{-1} \sinh(a_k \omega)} \frac{d^3 p}{(2\pi)^3}.$$

On the Brillouin zone $\hat{p} \geq 2|p|/\pi$ and $a_k^2 \hat{p}^2/2 \leq 6$; as $\operatorname{arccosh}(1+y)/\sqrt{2y}$ decreases on $(0, 6]$, $a_k \omega \geq \kappa a_k \hat{p}$ with $\kappa = \operatorname{arccosh} 7/\sqrt{12} > 0.76$, and $a_k^{-1} \sinh(a_k \omega) \geq \omega$. Since $|\hat{h}_k|$ is bounded uniformly in k , the integrands are dominated by a multiple of $e^{-8t_f |p|^2/\pi^2}/|p|$, which is integrable on $\mathbb{R}^3$, and they converge pointwise to $|\hat{h}(p)|^2 e^{-2t_f |p|^2} e^{-r|p|}/(2|p|)$. By dominated convergence $D_k(r')/D_k(r)$ tends to the corresponding positive ratio of continuum integrals; for $h = e^{-|x|^2}$, $t_f = 0.05$, $r = \frac{1}{2}$ and $r' = 1$ it is 0.6337.

Remark 53. The inequality also holds along the right tuning in two-dimensional $SU(N)$ lattice gauge theory, which is interacting and nonabelian but has no propagating gluons. Take the lattice $a_k \mathbb{Z} \times a_k \mathbb{Z}_{L_k}$ with $a_k L_k = \ell$ fixed, the weight $\prod_p \exp(\frac{\beta_k}{N} \operatorname{Re} \operatorname{Tr} U_p)$ with $\beta_k = 2N/(g^2 a_k^2)$, and the temporal gauge. A gauge-invariant function of the links of one slice depends only on the conjugacy class of their holonomy V around the circle; this includes every spatially flowed and spatially smeared slice field, since the flow is gauge covariant. Let w_β be the probability density proportional to $e^{\frac{\beta}{N} \operatorname{Re} \operatorname{Tr} U}$ with respect to Haar measure and, for an irreducible representation λ of dimension d_λ and character χ_λ , $\hat{c}_\lambda(\beta) = d_\lambda^{-1} \int \chi_\lambda w_\beta dU$. Between two slices put $W_x = V_x' V_x^{-1}$; the W_x are independent with density w_{β_k} , which is invariant under conjugation and inversion, and

$$V_1' \cdots V_L' = W_1 (V_1 W_2 V_1^{-1}) \cdots \\ \cdots (V_1 \cdots V_{L-1} W_L (V_1 \cdots V_{L-1})^{-1}) V_1 \cdots V_L,$$

so the new holonomy is QV with Q distributed as the L_k -fold convolution of w_{β_k} , independent of V . By Schur's lemma (Lemma 120) the transfer matrix acts on class functions by $\chi_\lambda \mapsto \hat{c}_\lambda(\beta_k)^{L_k} \chi_\lambda$. The coefficients are real, since w_β is invariant under $U \mapsto \bar{U}$, nonnegative, since the exponential of $\frac{\beta}{2N}(\chi_F + \bar{\chi}_F)$ is a sum of characters with nonnegative coefficients, and below 1 for $\lambda \neq 0$. For a class function f with $\int f dU = 0$ and isotypic components of norms $|\hat{f}_\lambda|$, the slice correlation is therefore

$$D_k(r) = \sum_{\lambda \neq 0} |\hat{f}_\lambda|^2 e^{-r E_{\lambda,k}}, \quad E_{\lambda,k} = -\frac{\ell}{a_k^2} \log \hat{c}_\lambda(\beta_k).$$

Laplace's method¹⁶ at the unique maximum $U = 1$ of $\operatorname{Re} \operatorname{Tr} U$, with $U = e^X$, $\operatorname{Re} \operatorname{Tr} e^X = N + \frac{1}{2} \operatorname{Tr} X^2 + O(|X|^4)$, $\operatorname{Re} \operatorname{Tr} X^3 = 0$, $\operatorname{Tr} \rho_\lambda(X) = 0$ and $\sum_a \rho_\lambda(T_a)^2 = -2C_2(\lambda)$ for an orthonormal basis of $\mathfrak{su}(N)$ in the norm $-\operatorname{Tr} X^2$, gives $\hat{c}_\lambda(\beta) = 1 - NC_2(\lambda)/\beta + O(\beta^{-2})$, with the Casimir normalised by $C_2 = (N^2 - 1)/(2N)$ on the fundamental representation (Lemma 121). Hence $E_{\lambda,k} \rightarrow g^2 \ell C_2(\lambda)/2$, and since $D_k(r) \leq \|f\|^2$,

$$\liminf_k \frac{D_k(r')}{D_k(r)} \geq \frac{|\hat{f}_\mu|^2}{\|f\|^2} e^{-r' g^2 \ell C_2(\mu)/2} > 0$$

¹⁶ For a smooth function with a unique nondegenerate maximum, $\int e^{\beta \varphi} h$ has an expansion in powers of β^{-1} whose coefficients are Gaussian moments at the maximum; Lemma 121 carries it out here.

for every $\mu \neq 0$ with $\hat{f}_\mu \neq 0$; for $f = \text{Re Tr } V$ the ratio is exactly $e^{-(r'-r)E_{F,k}} \rightarrow e^{-(r'-r)g^2\ell(N^2-1)/(4N)}$. This is the exact solution of two-dimensional lattice gauge theory, whose continuum limit is the measure of [32, 40, 52], written as the inequality. In two dimensions g^2 has the dimension of a mass squared and there is no ultraviolet problem, so the argument has no four-dimensional analogue. Nor does the spatial flow give the inequality by itself: it damps the spectral weight at large spatial momentum and leaves the weight at zero spatial momentum unchanged. For the free lattice field of Remark 52 with a mass M_k and the same flow, every energy is at least $m_k = a_k^{-1} \text{arccosh}(1 + a_k^2 M_k^2/2)$, so $D_k(r') \leq e^{-m_k(r'-r)} D_k(r)$, which tends to zero relative to $D_k(r)$ when $a_k M_k$ is fixed.

A route to the slice ratio in four dimensions replaces the packet by its conditional expectations given averages of the field over blocks of increasing size and controls the error of each step. The free field shows in which norm such errors can be small. In the ordinary L^2 norm the first step already fails: conditioning on the averages over blocks of side two leaves a fixed fraction of the variance of an energy packet, at every lattice spacing.

Proposition 54. *Let $d \geq 2$, let Λ be a periodic box in $\mathbb{Z}^d$ with even sides, let $\mu > 0$, and let ϕ be the Gaussian field on Λ with density proportional to $\exp(-\frac{1}{2} \sum_{x,v} (\nabla_v \phi(x))^2 - \frac{\mu^2}{2} \sum_x \phi(x)^2)$, where $\nabla_v \phi(x) = \phi(x + e_v) - \phi(x)$. Let $\mathcal{Z}$ be a set of blocks $B = 2z + \{0, 1\}^d$, let $\mathfrak{B}$ be the σ -algebra generated by the averages of ϕ over the blocks in $\mathcal{Z}$, let $f \geq 0$ on Λ , and let*

$$F = \sum_{x \in \Lambda} f(x) \sum_{v=1}^d :(\nabla_v \phi(x))^2:,$$

where $:X^2: = X^2 - \mathbb{E}X^2$. For $B = 2z + \{0, 1\}^d$ put $x_B = 2z$, $y_B = 2z + e_1 + e_2$ and

$$\alpha_B = \frac{1}{4} \sum_{w \in \{x_B, y_B\}} \left(df(w) + \sum_{v=1}^d f(w - e_v) \right),$$

and $\sigma^2 = 2/(2d + \mu^2)$. (i) $\mathbb{E} \text{Var}(F | \mathfrak{B}) \geq 2\sigma^4 \sum_{B \in \mathcal{Z}} \alpha_B^2$.

(ii) $\text{Var} F \leq 2d^2 \sum_x f(x)^2$.

(iii) Let $g \geq 0$ be a Lipschitz function on $\mathbb{R}^d$ with compact support and $\int g^2 > 0$. For $n \geq 1$ let Λ_n contain $n \text{supp } g$ in a fundamental domain, $f_n(x) = g(x/n)$, let $\mathcal{Z}_n$ contain every block that meets $n \text{supp } g$, and let $\mu = \mu_n \rightarrow 0$. Then for the packets F_n of f_n ,

$$\liminf_{n \rightarrow \infty} \frac{\mathbb{E} \text{Var}(F_n | \mathfrak{B}_n)}{n^d \int g^2} \geq 2^{1-d}, \quad \limsup_{n \rightarrow \infty} \frac{\text{Var} F_n}{n^d \int g^2} \leq 2d^2,$$

$$so \mathbb{E} \text{Var}(F_n | \mathfrak{B}_n) \geq (2^{-d} d^{-2} - o(1)) \text{Var} F_n.$$

Proof. (i) No two of the sites x_B, y_B , $B \in \mathcal{Z}$, are neighbours: their coordinate sums are even, and a neighbour of a site has a coordinate sum of the other parity. Let $\mathfrak{H} \supseteq \mathfrak{B}$ be generated by $\mathfrak{B}$ and by $\phi(w)$ for every w outside $\{x_B, y_B : B \in \mathcal{Z}\}$. By the law of total variance, $\text{Var}(F | \mathfrak{B}) = \mathbb{E}[\text{Var}(F | \mathfrak{H}) | \mathfrak{B}] + \text{Var}(\mathbb{E}[F | \mathfrak{H}] | \mathfrak{B})$, so $\mathbb{E} \text{Var}(F | \mathfrak{B}) \geq \mathbb{E} \text{Var}(F | \mathfrak{H})$. In the

exponent of the density, a site w enters through $(d + \frac{\mu^2}{2})\phi(w)^2$ and through $-\phi(w)\phi(w')$ for each of its $2d$ neighbours w' . Given $\mathfrak{H}$, the block average fixes $s_B = \phi(x_B) + \phi(y_B)$, and with $u_B = \phi(x_B) - \phi(y_B)$ we have $\phi(x_B)^2 + \phi(y_B)^2 = \frac{1}{2}(s_B^2 + u_B^2)$. Since the free sites have only $\mathfrak{H}$ -measurable neighbours, the conditional density of $(u_B)_{B \in \mathcal{Z}}$ is proportional to $\prod_B \exp(-\frac{1}{4}(2d + \mu^2)u_B^2 + c_B u_B)$ with $\mathfrak{H}$ -measurable c_B : the u_B are conditionally independent Gaussians of variance σ^2 . A term $(\nabla_v \phi(w))^2$ contains at most one free site, and if it contains x_B , then $\phi(x_B) = \frac{1}{2}(s_B + u_B)$ shows that it is $\frac{1}{4}u_B^2$ plus a term linear in u_B plus an $\mathfrak{H}$ -measurable term; the same holds for y_B with $-u_B$. The terms containing x_B are those with $w = x_B$ and those with $w = x_B - e_v$, d of each kind, so

$$F = \sum_{B \in \mathcal{Z}} (\alpha_B u_B^2 + \beta_B u_B) + F_0$$

with $\mathfrak{H}$ -measurable β_B and F_0 . Writing $u_B = m_B + v_B$ with m_B the conditional mean, $\alpha_B u_B^2 + \beta_B u_B = \alpha_B v_B^2 + (2\alpha_B m_B + \beta_B)v_B + \text{const}$, and for a centred Gaussian v of variance σ^2 , $\text{Var}(\alpha v^2 + \gamma v) = 2\alpha^2 \sigma^4 + \gamma^2 \sigma^2$. By conditional independence, $\text{Var}(F | \mathfrak{H}) \geq 2\sigma^4 \sum_B \alpha_B^2$.

(ii) For centred jointly Gaussian X, Y , $\text{Cov}(:X^2:, :Y^2:) = 2\text{Cov}(X, Y)^2$. With $G_{v\rho}(x, y) = \mathbb{E}[\nabla_v \phi(x) \nabla_\rho \phi(y)]$,

$$\begin{aligned} \text{Var} F &= 2 \sum_{x,y} f(x)f(y) \sum_{v,\rho} G_{v\rho}(x,y)^2 \\ &\leq \sum_{x,y} (f(x)^2 + f(y)^2) \sum_{v,\rho} G_{v\rho}(x,y)^2. \end{aligned}$$

On the dual torus Λ^* , with $\hat{\rho}^2 = \sum_v |e^{ipv} - 1|^2$,

$$G_{v\rho}(x, y) = \frac{1}{|\Lambda|} \sum_{p \in \Lambda^*} \frac{(e^{ipv} - 1)(e^{-ip\rho} - 1)}{\hat{\rho}^2 + \mu^2} e^{ip \cdot (x-y)},$$

and by Parseval $\sum_y G_{v\rho}(x, y)^2 = |\Lambda|^{-1} \sum_p |e^{ipv} - 1|^2 |e^{ip\rho} - 1|^2 (\hat{\rho}^2 + \mu^2)^{-2} \leq 1$, since each factor $|e^{ipv} - 1|^2$ is at most $\hat{\rho}^2$. The same holds for the sum over x , and (ii) follows.

(iii) If $\text{Lip } g = K$, every site entering α_B lies within distance 3 of $2z$, so $|\alpha_B - dg(2z/n)| \leq 3dK/n$. Hence $\sum_{B \in \mathcal{Z}_n} \alpha_B^2 \geq \sum_z (dg(2z/n) - 3dK/n)^2$, the sum running over z with $2z \in n \text{supp } g$, and this is $d^2(n/2)^d \int g^2(1 + o(1))$ as a Riemann sum on the grid $(2/n)\mathbb{Z}^d$. Since $\sigma^4 = 4/(2d + \mu_n^2)^2 \rightarrow d^{-2}$, (i) gives the first bound. By (ii) and $\sum_x g(x/n)^2 = n^d \int g^2(1 + o(1))$, the second follows. $\square$

For $d = 4$ and the energy density, $\text{Var} F_n$ grows like $(\ell/a)^4$ at a fixed physical size ℓ of the packet, and no step of conditioning on block averages removes more than a fraction $1 - 2^{-d}d^{-2} + o(1)$ of it. The reflected norm $C_k(0) = \omega_k(\vartheta(\tilde{A}_k)A_k)$, which renormalises the packet in Theorem 48, involves the covariance only between the support of ϑf_0 and that of f_0 , at a fixed physical distance. The errors of a blocking route have therefore to be measured in that norm; Figure 12 shows the blocks between the reflection plane and the packet.

(a) holonomy on the circle

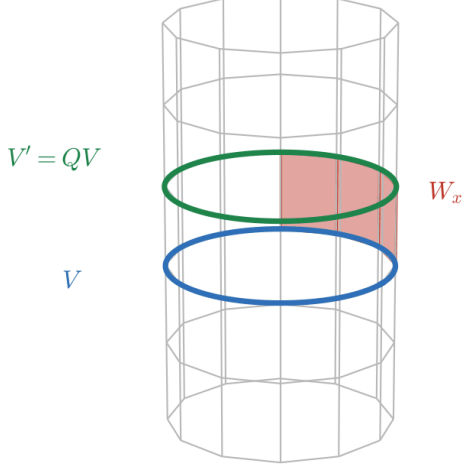
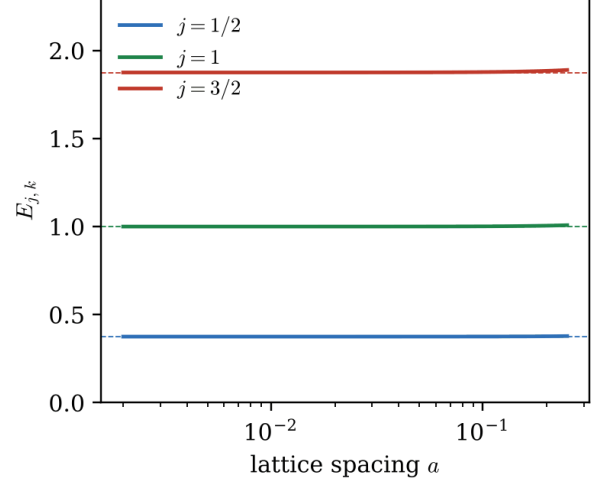
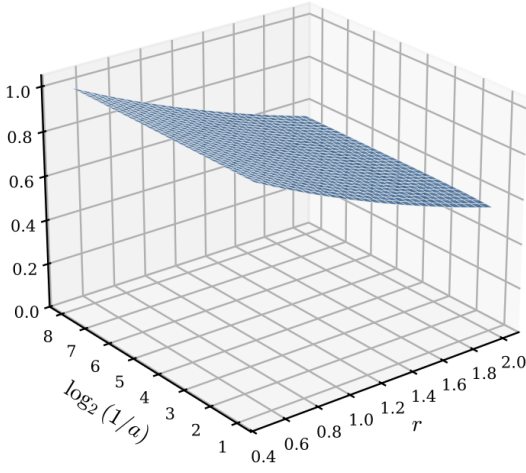
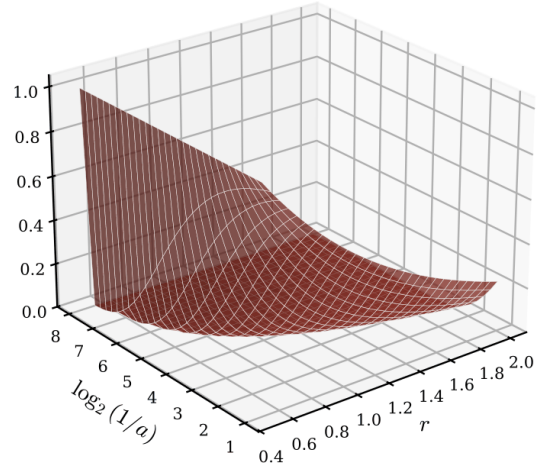
(b) $SU(2)$, $g^2 = \ell = 1$; limits $j(j+1)/2$ 

FIG. 10. Two-dimensional $SU(N)$ (Remark 53). (a) The lattice $a_k\mathbb{Z} \times a_k\mathbb{Z}_{L_k}$ drawn as a cylinder: the holonomy V of one slice, the holonomy $V' = QV$ of the next, and the plaquette variables W_x between them. (b) The energies $E_{j,k} = -(\ell/a_k^2) \log \hat{c}_j(\beta_k)$ of $SU(2)$ for $g^2 = \ell = 1$, with $\hat{c}_j(\beta) = I_{2j+1}(\beta)/I_1(\beta)$, against the lattice spacing; the dashed lines are the limits $j(j+1)/2$.

(a) $SU(2)$ in two dimensions

(b) mass 1/2 in lattice units



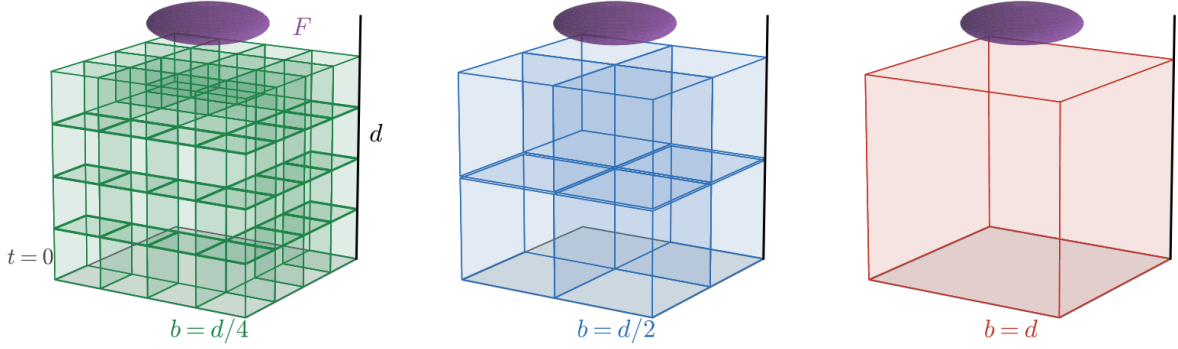
height: $D_k(r)/D_k(1/2)$ for $g^2 = \ell = 1$

FIG. 11. The ratio $D_k(r)/D_k(\frac{1}{2})$ over the separation r and the lattice spacing a_k . (a) Two-dimensional $SU(2)$ with $f = \text{Re Tr } V$ and $g^2 = \ell = 1$, where the ratio is $e^{-(r-1/2)E_{1/2,k}}$ and converges as $a_k \rightarrow 0$. (b) The term $e^{-(r-1/2)m_k}$ of the lowest energy m_k of a free field with mass $\frac{1}{2}$ in lattice units, which bounds the ratio from above and tends to zero at every $r > \frac{1}{2}$.

E. Higher correlations

Theorem 48 controls two renormalised fields at a time, and part (i) of it gives nothing for more: Z_k may diverge while the fields $\Phi_x \circ \Phi_{s_k}$ stay bounded, so a bound on the norm of the renormalised packet as an operator is not available. A reflection positive process with a fixed gap and a two-point function independent of the cutoff can have divergent four-

point functions at separated times (Remark 58), so no statement about the two-point function alone can replace the following hypothesis on the truncated correlations. In this subsection Φ is real-valued, as the energy of the plaquettes is; for a complex field the conjugate field has to be adjoined and (T1) imposed on all mixed cumulants. Write $\varphi_k(x) = \Phi_x \circ \Phi_{s_k}$, ω_k^T for the truncated correlations (joint cumulants) under ω_k , and for $f \in$



block averages of side b between the plane $t = 0$ and the packet F at height d

FIG. 12. Block averages between the reflection plane $t = 0$ and a packet F at physical height d , for blocks of side $d/4$, $d/2$ and d . By Proposition 54 the first conditioning step already leaves a fixed fraction of the ordinary variance of the packet, so the errors of such steps have to be measured in the reflected norm.

$C_c(\mathbb{R}^4)$

$$A_k(f) = \sum_{x \in \mathbb{Z}^4} a_k^4 f(a_k x) (\varphi_k(x) - \omega_k(\varphi_k(x))),$$

so that $A_k = A_k(f_0)$ and $\bar{A}_k(f) = A_k(\bar{f})$. Let $w_k : \mathbb{Z}^4 \rightarrow [0, \infty)$ be even functions with $W_k = \sum_x w_k(x) < \infty$, let $r_k > 0$ be amplitude scales, and consider:

(T1) *Tree bounds.* For every $n \geq 2$ there is K_n , independent of k , with

$$|\omega_k^T(\varphi_k(x_1); \dots; \varphi_k(x_n))| \leq K_n r_k^n \sum_T \prod_{\{i,j\} \in T} w_k(x_i - x_j)$$

for all $x_1, \dots, x_n \in \mathbb{Z}^4$, the sum running over the trees T on $\{1, \dots, n\}$.

(T2) *Saturation.* $C_k(0) \geq c r_k^2 a_k^4 W_k$ for a $c > 0$ and every k .

(T3) *Physical volume.* $B_* = \sup_k a_k^4 W_k < \infty$.

The amplitude r_k separates the size of the field from the volume over which it is correlated; Remark 57 shows why it is needed. For a compact $Q \subset \mathbb{R}^4$ let $|Q|_k = a_k^4 \# \{x \in \mathbb{Z}^4 : a_k x \in Q\}$; for a cube of side L it is at most $(L + a_0)^4$.

Lemma 55. Under (T1), for $n \geq 2$ and bounded $f_1, \dots, f_n$ with compact supports, f_1 supported in Q ,

$$|\omega_k^T(A_k(f_1); \dots; A_k(f_n))| \leq K_n n^{n-2} r_k^n |Q|_k (a_k^4 W_k)^{n-1} \prod_{j=1}^n \|f_j\|_\infty.$$

Proof. Joint cumulants of order $n \geq 2$ are multilinear and do not change when constants are added to the variables, so the

left side is at most $a_k^{4n} \sum_{x_1, \dots, x_n} \prod_j |f_j(a_k x_j)|$ times the absolute value of the truncated correlation, the sums being finite. Insert (T1) and fix a tree T . With x_1 fixed, sum over the other points one at a time, each time over a leaf of what remains of T other than 1; each such sum is at most W_k . The remaining sum over x_1 has at most $|Q|_k a_k^{-4}$ nonzero terms. There are n^{n-2} trees. $\square$

Theorem 56. Assume (T1)–(T3).

(i) For $n \geq 2$, f_j as in the lemma,

$$|\omega_k^T(Z_k A_k(f_1); \dots; Z_k A_k(f_n))| \leq K_n n^{n-2} c^{-n/2} B_*^{n/2-1} |Q|_k \prod_{j=1}^n \|f_j\|_\infty,$$

and the moments $\omega_k(\prod_j Z_k A_k(f_j))$ with all f_j supported in a cube Q are bounded by $M_n(Q) \prod_j \|f_j\|_\infty$, with $M_n(Q)$ independent of k .

(ii) Every sequence of k has a subsequence along which

$$S_n(f_1, \dots, f_n) = \lim_k \omega_k \left(\prod_{j=1}^n Z_k A_k(f_j) \right)$$

exists for all n and all $f_j \in C_c(\mathbb{R}^4)$. Each S_n is symmetric, multilinear, bounded by $M_n(Q) \prod_j \|f_j\|_\infty$ on functions supported in Q , hence a distribution on $(\mathbb{R}^4)^n$; it is invariant under all translations of $\mathbb{R}^4$ if the ω_k are translation invariant, and the kernel K of Theorem 48 along the subsequence is $K(y, y') = S_2(\bar{f}_0(\cdot - y) \circ \theta, \bar{f}_0(\cdot - y'))$, with θ as in (iii).

(iii) Put $S_0 = 1$, $S_1 = 0$, and let θ be the reflection $t \mapsto -t$ of $\mathbb{R}^4$. For every finite family of $c_\alpha \in \mathbb{C}$ and $f_{\alpha,j} \in C_c(\mathbb{R}^4)$ supported in $\{t > 0\}$,

$$\sum_{\alpha, \beta} \bar{c}_\alpha c_\beta S \left(\prod_j \overline{f_{\alpha,j}} \circ \theta \cdot \prod_j f_{\beta,j} \right) \geq 0,$$

where S of a product of n test functions means S_n . Hence the Osterwalder–Schrader construction gives a Hilbert space $\mathcal{H}$ with vacuum Ω and a dense set of vectors $\Psi(P)$ indexed by polynomials P in the smeared fields at positive times, with $\langle \Psi(P), \Psi(P') \rangle = S(\theta \bar{P} \cdot P')$; in particular the separated higher correlations of the renormalised fields are finite. The vector $\Psi(P_0)$ of the field smeared by f_0 is a unit vector orthogonal to Ω .

(iv) Suppose the ω_k are translation invariant. Then the time translations by $t \geq 0$ define on $\mathcal{H}$ a strongly continuous semigroup e^{-tH} with $H \geq 0$ self-adjoint, and the spatial translations a strongly continuous unitary group commuting with H . Moreover there is a dyadic $\sigma > 0$ with $C_k(\sigma) \geq \frac{1}{2}C_k(0)$ for all large k , so the limit of Theorem 48 is nontrivial.

(v) If $a_k^4 W_k \rightarrow 0$ along the subsequence, every S_n is given by Wick's rule from S_2 : the limit is the Gaussian field with covariance S_2 .

(vi) If the ω_k are translation invariant and the hypothesis of Theorem 45 holds, then 0 is a simple eigenvalue of H and $\sigma(H) \subset \{0\} \cup [m_*, \infty)$ with $m_* = \log(1/\alpha)/(2\ell)$, and $\Omega^\perp \neq 0$.

Proof. (i) By (T2), $Z_k^n = C_k(0)^{-n/2}$ is at most $(c r_k^2 a_k^4 W_k)^{-n/2}$, multiply the lemma by it, which cancels r_k^n , and use $(a_k^4 W_k)^{n/2-1} \leq B_*^{n/2-1}$ by (T3) for $n \geq 2$. A moment is the sum over partitions of $\{1, \dots, n\}$ of the products of the cumulants of the blocks; blocks of size one vanish because $A_k(f)$ is centred, and each block contributes the bound just proved with $|Q|_k \leq (L + a_0)^4$ for a cube of side L . This gives $M_n(Q)$.

(ii) For each cube $Q_L = [-L, L]^4$, $L \in \mathbb{N}$, the continuous functions supported in Q_L form a separable space under the supremum norm. Choose a countable dense set $\mathcal{F}$ of functions; by (i) the countably many sequences $\omega_k(\prod Z_k A_k(f_j))$, $f_j \in \mathcal{F}$, are bounded, and a diagonal subsequence makes all of them converge. By multilinearity,

$$\begin{aligned} & \left| \omega_k \left(\prod_j Z_k A_k(f_j) \right) - \omega_k \left(\prod_j Z_k A_k(g_j) \right) \right| \\ & \leq M_n(Q) \sum_i \|f_i - g_i\|_\infty \prod_{j \neq i} \max(\|f_j\|_\infty, \|g_j\|_\infty) \end{aligned}$$

for functions supported in Q , so the limits exist for all $f_j \in C_c(\mathbb{R}^4)$ and obey the same bound. Symmetry and multilinearity pass to the limit. A multilinear form bounded in this way is separately continuous on $C_c^\infty(\mathbb{R}^4)$, so the Schwartz kernel theorem makes it a distribution. For dyadic y and large k , $y/a_k \in \mathbb{Z}^4$ and translation invariance of ω_k and of the flow (Proposition 47(i)) give $\omega_k(\prod Z_k A_k(f_j(\cdot - y))) = \omega_k(\prod Z_k A_k(f_j))$. For a real y take dyadic $y_m \rightarrow y$; the functions $f_j(\cdot - y_m)$ converge uniformly to $f_j(\cdot - y)$ with supports in one cube, so the bound above carries the invariance to y . The formula for K follows from $A_k^y = A_k(f_0(\cdot - y))$ and the argument of (iii) with $n = 2$.

(iii) Let ϑ be the lattice reflection of Appendix F and θ_k the corresponding reflection of $a_k \mathbb{Z}^4$; it differs from θ by a translation of length at most a_k . Since Φ is real and ϑ commutes with the flow, $\vartheta(\bar{A}_k(f)) = A_k(\bar{f} \circ \theta_k)$. Let $P_k = \sum_\beta c_\beta \prod_j Z_k A_k(f_{\beta,j})$ and $R_k = \sum_\alpha \bar{c}_\alpha \prod_j Z_k A_k(\bar{f}_{\alpha,j} \circ \theta)$, so that $\vartheta(\bar{P}_k)$ is R_k with θ replaced by θ_k . Reflection positivity gives $\omega_k(\vartheta(\bar{P}_k)P_k) \geq 0$. By (i), applied in a cube containing

all the supports and their reflections, and by multilinearity, $|\omega_k(\vartheta(\bar{P}_k)P_k) - \omega_k(R_k P_k)|$ is at most a constant independent of k times $\max_{\alpha,j} \|f_{\alpha,j} \circ \theta_k - f_{\alpha,j} \circ \theta\|_\infty$, which tends to zero by uniform continuity, while $\omega_k(R_k P_k)$ tends to the left side of (iii) by (ii). So the limit of the nonnegative numbers is the left side of (iii). The construction of $\mathcal{H}$ is the usual quotient of the polynomials at positive times by the null space of this form [62]. By the same approximation of θ by θ_k , $\|\Psi(P_0)\|^2 = \lim_k Z_k^2 C_k(0) = 1$, and $\langle \Omega, \Psi(P_0) \rangle = S_1(f_0) = 0$.

(iv) For a polynomial P at positive times let $\tau_t P$ be its time translate by $t \geq 0$, again at positive times. By translation invariance and $\theta \circ \tau_t = \tau_{-t} \circ \theta$, $\langle \Psi(\tau_t P), \Psi(P') \rangle = \langle \Psi(P), \Psi(\tau_t P') \rangle$, and $q_P(t) = \|\Psi(\tau_t P)\|^2 = \langle \Psi(P), \Psi(\tau_{2t} P) \rangle$. If P has degree at most d , every partition in (i) of the products in $q_P(t)$ has at most d blocks, each bounded by a multiple of $|Q|_k$ for a cube Q of side $O(1+t)$, so $q_P(t) \leq C_P(1+t)^{4d}$. The Cauchy–Schwarz inequality gives $q_P(t) \leq q_P(0)^{1/2} q_P(2t)^{1/2}$, hence by induction $q_P(t) \leq q_P(0)^{1-2^{-m}} q_P(2^m t)^{2^{-m}}$ for every m , and letting $m \rightarrow \infty$, $q_P(t) \leq q_P(0)$. So τ_t maps null vectors to null vectors and extends to a symmetric contraction $T(t)$ of $\mathcal{H}$; $T(s)T(t) = T(s+t)$, and $T(t) = T(t/2)^2 \geq 0$. By the bound of (ii) and uniform continuity of the translated test functions, $\|\Psi(\tau_t P) - \Psi(P)\| \rightarrow 0$ as $t \rightarrow 0$, so the semigroup is strongly continuous on a dense set, hence everywhere, and $T(t) = e^{-tH}$ with $H \geq 0$ self-adjoint [62]. Spatial translations preserve the form, have the opposite translations as inverses, commute with τ_t , and are strongly continuous by the same argument.

For the ratio let $f_{0,t} = f_0(\cdot - (t, 0))$. For dyadic $t > 0$ and large k , translation invariance gives $\tau_{t/a_k} A_k = A_k(f_{0,t})$, so by (iii)

$$C_k(t) - C_k(0) = \omega_k(A_k(\bar{f}_0 \circ \theta_k) A_k(f_{0,t} - f_0)),$$

a covariance of two centred fields. The case $n = 2$ of (i) bounds Z_k^2 times it by $L_0 \|f_{0,t} - f_0\|_\infty$ with $L_0 = (K_2/c) \sup_k |Q|_k \|f_0\|_\infty$, Q a cube containing the supports of $f_0 \circ \theta_k$ for all k . By uniform continuity of f_0 there is a dyadic $\sigma > 0$ with $L_0 \|f_{0,\sigma} - f_0\|_\infty \leq \frac{1}{2}$; then $C_k(\sigma)/C_k(0) \geq \frac{1}{2}$, and Theorem 48(iv) applies.

(v) For $n \geq 3$ the exponent $n/2 - 1$ in (i) is positive, so with $(a_k^4 W_k)^{n/2-1}$ in place of $B_*^{n/2-1}$ the cumulants of order $n \geq 3$ of the renormalised fields tend to zero. The moments are the sums over partitions of products of cumulants, so in the limit only the partitions into pairs remain, which is Wick's rule.

(vi) As in the proof of Theorem 45, $\mathbb{T}_k$ has gap at least $\log(1/\alpha)/(2n_k)$ on the complement of Ω_k . For a polynomial P at positive times let $v_k = \Psi_{P_k} - \omega_k(P_k)\Omega_k$ with P_k its lattice version. For dyadic t and large k ,

$$\omega_k(\vartheta(\bar{P}_k)\tau_{t/a_k} P_k) - |\omega_k(P_k)|^2 = \langle v_k, \mathbb{T}_k^{t/a_k} v_k \rangle \leq e^{-m_* t} \|v_k\|^2,$$

and both sides are finite combinations of moments, which converge by (ii) and the approximation of θ in (iii). The limit is $\langle v, e^{-tH} v \rangle \leq e^{-m_* t} \|v\|^2$ for $v = \Psi(P) - \langle \Omega, \Psi(P) \rangle \Omega$. These v are dense in $\Omega^\perp$, so the inequality holds on $\Omega^\perp$ for all $t \geq 0$, and the spectral theorem gives the spectral statement as in Theorem 45. By (iii), $\Psi(P_0)$ is a unit vector in $\Omega^\perp$. $\square$

Part (iv) shows that the order-two cases of (T1) and (T2) already contain the lower bound on the ratio that Theorem 48(iv)

needs; they are not harmless combinatorial assumptions, and proving them along the continuum tuning is at least as hard as that lower bound. Part (vi) passes the conditional mass gap of Theorem 45 to every polynomial in the smeared fields. Neither part proves Euclidean invariance, independence of the limit from the subsequence, or the reconstruction of the field operators, for which the growth of $M_n(Q)$ in n would also be needed.

Remark 57. The amplitude scale cannot be dropped for a bounded field. Suppose $\text{Var } \varphi_k(0) \rightarrow 0$, as for the energy of the flowed plaquettes in lattice units when the flowed links approach the identity. Reflection invariance of ω_k , the Cauchy–Schwarz inequality and stationarity give

$$C_k(0) \leq \omega_k(|A_k|^2) \leq \left(a_k^4 \sum_x |f_0(a_k x)| \right)^2 \text{Var } \varphi_k(0) \rightarrow 0.$$

With $r_k = 1$, (T2) then forces $a_k^4 W_k \rightarrow 0$, and by part (v) the limit is Gaussian. With r_k^2 of the order of $\text{Var } \varphi_k(0)$ this is no longer forced. Choosing r_k does not prove (T1) or (T2).

Remark 58. Let $0 < \varepsilon < 1$, $m > 0$, let π be the probability measure on $\{-1, 0, 1\}$ with weights $\varepsilon/2$, $1 - \varepsilon$, $\varepsilon/2$, and let $P_t = e^{-mt}I + (1 - e^{-mt})\Pi$ on $L^2(\pi)$, where $\Pi f = \pi(f)$ is the projection onto the constants. This is a symmetric Markov semigroup with $P_t \geq 0$ and gap m , so its stationary process is reflection positive. For $X(s) = s/\sqrt{\varepsilon}$ we have $\pi(X) = \pi(X^3) = 0$, $\pi(X^2) = 1$, $\pi(X^4) = 1/\varepsilon$ and $P_t X = e^{-mt}X$, so the two-point function is e^{-mt} for every ε , while for $t_1 \leq t_2 \leq t_3 \leq t_4$ and $q_i = e^{-m(t_{i+1} - t_i)}$ the four-point function is $q_3 \pi(X P_{t_2 - t_1}(X P_{t_3 - t_2} X^2)) = q_1 q_3 (q_2/\varepsilon + 1 - q_2)$, with truncated part $q_1 q_3 (q_2/\varepsilon - q_2 - 2q_2^2)$; both diverge as $\varepsilon \rightarrow 0$. By part (i), no choice of amplitude scales makes (T1)–(T3) hold for this family with constants independent of ε .

The hypotheses (T1)–(T3) are uniform cluster estimates for the continuum tuning $\beta_k \rightarrow \infty$, and the next proposition shows that they are estimates at a critical point: with the order-two cases alone the lattice gap is at most a fixed multiple of a_k , and the tree majorant carries a fixed fraction of its weight at a fixed physical distance. Let $m_k \in [0, \infty]$ be the gap of $\mathbb{T}_k$ on the complement of Ω_k , so that $\|\mathbb{T}_k v\| \leq e^{-m_k} \|v\|$ for $v \perp \Omega_k$, and let $t_0 > 0$ be the least time in the support of f_0 .

Proposition 59. (i) *For every dyadic $\sigma > 0$ and every large k , $C_k(\sigma) \leq e^{-m_k \sigma / a_k} C_k(0)$. If $m_k / a_k \rightarrow \infty$, every limit of Theorem 48 is trivial; this is the case for $\beta_k = \beta$ fixed in the range of Theorem 5, where $m_k \geq m_G(\beta)$.*

(ii) *Suppose the ω_k are translation invariant and (T1) for $n = 2$ and (T2) hold. Then $m_k \leq a_k \log 2 / \sigma$ for every large k , where σ is the dyadic number of Theorem 56(iv), which depends only on K_2/c and f_0 . If also the hypothesis of Theorem 45 holds, then $m_* a_k \leq m_k \leq a_k \log 2 / \sigma$, and in particular $m_* \leq \log 2 / \sigma$.*

(iii) *Under (T1) for $n = 2$ and (T2), let Q be a cube containing the support of f_0 and $\gamma = c / (K_2 \|f_0\|_\infty^2 \sup_k |Q|_k)$. Then for every k*

$$\sum_{z \in \mathbb{Z}^4, |z_0| \geq R_k} w_k(z) \geq \gamma W_k, \quad R_k = \frac{2t_0}{a_k} - 1.$$

If $w_k(z) = e^{-\mu_k |z_0|} u_k(\bar{z})$ with $\bar{z}$ the spatial part of z and $u_k \geq 0$ summable on $\mathbb{Z}^3$, then $\mu_k R_k \leq \log(2/\gamma)$, so the physical time decay rate of the majorant is at most $\log(2/\gamma)/(2t_0 - a_k)$.

Proof. (i) In the proof of Theorem 48, $C_k(\sigma)/C_k(0) = \langle \Psi_k, \mathbb{T}_k^{\sigma/a_k} \Psi_k \rangle$ with Ψ_k a unit vector orthogonal to Ω_k . The operator $\mathbb{T}_k$ is self-adjoint and fixes Ω_k , so it preserves the complement, and the inequality follows. If $m_k/a_k \rightarrow \infty$ then $\rho(\sigma) = 0$ for every dyadic $\sigma > 0$ along every subsequence, and the limit is trivial by Theorem 48(iv). At fixed β in the range of Theorem 5 the gap is at least $m_G(\beta) > 0$ in lattice units, and $a_k \rightarrow 0$; the spatial flow does not change the transfer matrix (Proposition 47(iv)).

(ii) Theorem 56(iv) uses (T1) and (T2) only for $n = 2$, since the exponent $n/2 - 1$ in part (i) vanishes there, and gives $C_k(\sigma) \geq \frac{1}{2} C_k(0)$ for large k . With (i), $e^{-m_k \sigma / a_k} \geq \frac{1}{2}$. Under the hypothesis of Theorem 45, the proof of that theorem gives $m_k \geq \log(1/\alpha)/(2n_k) \geq m_* a_k$.

(iii) $C_k(0) = \omega_k(A_k(\bar{f}_0 \circ \theta_k) A_k(f_0))$, a covariance of two centred fields, by the proof of Theorem 56(iv). A point x in the support of the first sum has $a_k x_0 \leq -t_0 + a_k$, and a point y in the support of the second has $a_k y_0 \geq t_0$, so $|x_0 - y_0| \geq R_k$. By (T1) for $n = 2$,

$$\begin{aligned} C_k(0) &\leq K_2 r_k^2 \|f_0\|_\infty^2 a_k^8 \sum_{a_k y \in Q} \sum_{|x_0 - y_0| \geq R_k} w_k(x - y) \\ &\leq K_2 r_k^2 \|f_0\|_\infty^2 |Q| a_k^4 \sum_{|z_0| \geq R_k} w_k(z), \end{aligned}$$

and (T2) bounds the left side below by $c r_k^2 a_k^4 W_k$. For the product form let $R' = \lceil R_k \rceil$; if $R' \leq 0$ there is nothing to prove, and otherwise

$$\frac{\sum_{|j| \geq R'} e^{-\mu_k |j|}}{\sum_{j \in \mathbb{Z}} e^{-\mu_k |j|}} = \frac{2e^{-\mu_k R'}}{1 + e^{-\mu_k}} \leq 2e^{-\mu_k R_k},$$

while the factor $\sum_{\bar{z}} u_k(\bar{z})$ cancels. So $2e^{-\mu_k R_k} \geq \gamma$. $\square$

So reflection positivity and decay do not supply (T2): at fixed β in the range of Theorem 5 the measures are reflection positive with exponential clustering, and $C_k(\sigma)/C_k(0) \rightarrow 0$, so no choice of w_k , r_k and constants satisfies the order-two cases of (T1) and (T2) together there. For the same reason no estimate whose constants give a lattice gap bounded below, as a convergent strong coupling expansion does [64], can give the tree bounds along the tuning: (T1) and (T2) hold, if at all, only along couplings where the lattice gap closes at least as fast as a_k . A mass gap in physical units, which bounds the decay of the two-point function, does not give the constants K_n either. By Theorem 56(v) the limit can be non-Gaussian only if $\liminf_k a_k^4 W_k > 0$; for the product majorants of part (iii) the time factor of W_k is $\coth(\mu_k/2) \geq 2/\mu_k$, which is of order $1/a_k$ at least, so the time direction alone is compatible with this. At a fixed physical flow time the flowed fields are smooth on the scale $\sqrt{\tau}$; majorants of product form for them are not proved here.

The order-two cases of (T1) and (T2) say more about the reference vector than part (iv) of Theorem 56 records: for a smooth f_0 they bound every time difference of the packet,

which makes its energy distribution tight with finite moments and gives the ratio with an explicit rate. Recall the spectral measures μ_k on $[0, \infty]$ of the proof of Theorem 48, so that $C_k(\sigma)/C_k(0) = \int e^{-\sigma E} d\mu_k(E)$.

Proposition 60. *Suppose the ω_k are translation invariant, (T1) holds for $n = 2$ and (T2) holds, and f_0 has $m \geq 1$ continuous time derivatives. Let Q be a cube containing the supports of $f_0(\cdot - (s, 0))$ for $0 \leq s \leq ma_0$, and $B_m = (K_2/c) \sup_k \|Q\| \|\partial_t^m f_0\|_\infty^2$.*

(i) $\|(1 - \mathbb{T}_k)^m \Psi_k\|^2 \leq B_m a_k^{2m}$ for every k , where $\Psi_k = Z_k \Psi_{A_k}$.

(ii) $\int ((1 - e^{-a_k E})/a_k)^{2m} d\mu_k(E) \leq B_m$ and $\mu_k([\Lambda, \infty]) \leq B_m (a_k + \Lambda^{-1})^{2m}$ for $\Lambda > 0$. Every weak limit μ of the μ_k satisfies $\mu(\{\infty\}) = 0$ and $\int E^{2m} d\mu(E) \leq B_m$. If (T1)–(T3) hold, the unit vector $\Psi(P_0)$ of Theorem 56(iii)–(iv) lies in the domain of H^m , with $\|H^m \Psi(P_0)\|^2 \leq B_m$.

(iii) If $a_k \sqrt{B_1} \leq \frac{1}{2}$, then $C_k(\sigma) \geq e^{-2\sqrt{B_1}\sigma} C_k(0)$ for every $\sigma \in a_k \mathbb{Z}_{\geq 0}$, and the gap m_k of Proposition 59 satisfies $m_k \leq -\log(1 - a_k \sqrt{B_1}) \leq 2\sqrt{B_1} a_k$.

Proof. (i) Let $D_k f_0 = \sum_{j=0}^m (-1)^j \binom{m}{j} f_0(\cdot - (ja_k, 0))$. By translation invariance $\mathbb{T}_k^j \Psi_{A_k} = \Psi_{\tau_j A_k}$ and $\tau_j A_k = A_k(f_0(\cdot - (ja_k, 0)))$, so $(1 - \mathbb{T}_k)^m \Psi_{A_k} = \Psi_{A_k(D_k f_0)}$, and as in the proof of Theorem 56(iv) its squared norm is the covariance $\omega_k(A_k(\overline{D_k f_0} \circ \theta_k) A_k(D_k f_0))$. The function $D_k f_0$ is supported in Q and is an m -th difference with step a_k , so $\|D_k f_0\|_\infty \leq a_k^m \|\partial_t^m f_0\|_\infty$. Lemma 55 with $n = 2$ bounds the covariance by $K_2 r_k^2 \|Q\| a_k^4 W_k \|D_k f_0\|_\infty^2$, and (T2) bounds $C_k(0) = Z_k^{-2}$ below by $c r_k^2 a_k^4 W_k$.

(ii) By the spectral theorem the left side of (i) is $\int (1 - e^{-a_k E})^{2m} d\mu_k$, the integrand being 1 at $E = \infty$. For $E \geq \Lambda$, $(1 - e^{-a_k E})/a_k \geq (1 - e^{-a_k \Lambda})/a_k \geq \Lambda/(1 + a_k \Lambda)$, because $1 - e^{-x} \geq x/(1+x)$; this gives the tail bound. The set $(\Lambda, \infty]$ is open in $[0, \infty]$, so $\mu((\Lambda, \infty]) \leq \liminf_k \mu_k((\Lambda, \infty]) \leq B_m \Lambda^{-2m}$, and $\mu(\{\infty\}) = 0$. For a continuous χ with $0 \leq \chi \leq 1$ and support in $[0, L]$, the same inequality gives $E \leq (1 + a_k L)(1 - e^{-a_k E})/a_k$ on that support, so $\int \chi E^{2m} d\mu = \lim_k \int \chi E^{2m} d\mu_k \leq B_m$; letting χ increase to 1 gives the moment bound. The vector $\Psi(P_0)$ has $\langle \Psi(P_0), e^{-\sigma H} \Psi(P_0) \rangle = \lim_k C_k(\sigma)/C_k(0) = \int e^{-\sigma E} d\mu$ for dyadic $\sigma > 0$ along the subsequence, so its spectral measure for H is μ .

(iii) By (i) with $m = 1$ and the Cauchy–Schwarz inequality, $\langle \Psi_k, (1 - \mathbb{T}_k) \Psi_k \rangle \leq a_k \sqrt{B_1}$, which is the hypothesis of Proposition 49(ii) with $C = \sqrt{B_1}$. Since Ψ_k is a unit vector orthogonal to Ω_k , the same quantity is at least $1 - e^{-m_k}$. $\square$

Part (iii) makes Proposition 59(ii) explicit: the order-two tree and saturation bounds pin the lattice gap below $2\sqrt{B_1} a_k$, a physical mass of at most $2\sqrt{B_1}$, and give the ratio of Theorem 48(iv) at every separation. The open input of Theorem 48 is therefore implied by these two bounds for one smooth f_0 ; nothing here proves them for the Wilson action.

XI. THE OPEN INPUTS

This section states the two unproved inputs of Sec. X, the cube estimate and the limit, and measures the first: what can be controlled uniformly in the coupling, and what any proof of the cube estimate has to produce.

A. The two open statements

Theorem 45 reduces a continuum mass gap to two statements, neither of which is proved here or, to our knowledge, anywhere else.

(D) *The cube estimate.* There are $\alpha < 1$ and $\ell > 0$ such that $D(n, \alpha)$ holds with $na \leq \ell$ along the continuum tuning $\beta = \beta(a) \rightarrow \infty$, uniformly in a .

(E) *The limit.* The renormalised Schwinger functions converge along $a_k = 2^{-k} a_0$ to Schwinger functions that satisfy the Osterwalder–Schrader axioms and are not those of the vacuum alone.

(D) and (E) together give the mass gap of the continuum theory; the remaining Clay requirements, Euclidean invariance and the local structure of the limit [47], belong to (E). (D) is a statement about a single cube of fixed physical size and a single coupling at each spacing; (E) is the construction problem, on which Balaaban’s renormalisation group analysis [5–7] gives the ultraviolet part in finite volume.

Remark 61. The following count relates (D) to a correlation length. Suppose that the gauge-fixed field responds to a change of one link on the boundary of a cube through a Poisson kernel with decay length ξ in lattice units. Then the response summed over the cube is of order ξ , so K_c is of order $n^3 \xi$ and $D(n, \alpha)$ holds once $n \gtrsim \xi/\alpha$, that is, once $\ell \gtrsim \xi a/\alpha$, a fixed multiple of the physical correlation length. The hypothesis is that the gauge-fixed links decorrelate on the scale ξ , and it is a statement about the gauge, not about the gauge-invariant observables: the gauge here is built on straight lines, and in gauges of axial type the gauge-fixed variables can fluctuate on all scales while gauge-invariant observables cluster. Propositions 62 and 63 carry the count out exactly for Gaussian fields, and Proposition 63 shows that the quadratic approximation of the forest gauge fluctuates in this way. Where it does, (D) fails in this gauge and has to be posed for block-averaged variables, where it becomes part of the renormalisation group problem. Sec. XII A carries this out in the quadratic approximation.

The count of Remark 61 can be carried out exactly for a Gaussian field, the quadratic approximation of the gauge-fixed theory at weak coupling, and shows what (D) asks of the gauge-fixed variables.

Proposition 62. *Let $d \geq 1$, $b > 0$, $t \geq 0$, let $n \geq 2$ be even and $V = [0, n)^d \cap \mathbb{Z}^d$. Given $\eta \in \mathbb{R}^{\mathbb{Z}^d \setminus V}$, let $\mathcal{W}(\cdot | \eta)$ be the Gaussian law on $\mathbb{R}^V$ with density proportional to*

$$\exp\left(-\frac{b}{2} \left[\sum_{\{x,y\} \cap V \neq \emptyset} (\phi_x - \phi_y)^2 + t \sum_{x \in V} \phi_x^2 \right] \right),$$

the first sum over nearest-neighbour pairs, with $\phi = \eta$ outside V . Define $C_{V,j}$, the classes c (residues modulo two) and K_c as above, with $d_V(\phi, \phi') = \sum_{x \in V} |\phi_x - \phi'_x|$ and $\delta(\eta_j, \eta'_j) = |\eta_j - \eta'_j|$, and put $N_n = (n/2)^d$. Then $K_c = \alpha_n(t)N_n$ for every class c , where

$$\alpha_n(t) = 1 - \frac{t}{n^d} \mathbf{1}^\top (t + L_n)^{-1} \mathbf{1}$$

and L_n is the Dirichlet Laplacian of V . In particular:

- (i) for $t = 0$, $K_c = N_n$ for every n and every b , so $D(n, \alpha)$ fails for every $\alpha < 1$;
- (ii) for $t > 0$, $1 - \alpha_n(t) \geq (5/6)^d t n^2 / (t n^2 + 10d)$; for $d = 4$, $t = m^2 a^2$ and $\ell/2 \leq na \leq \ell$ this gives $D(n, \alpha)$ with $\alpha = 1 - (5/6)^4 (m\ell)^2 / ((m\ell)^2 + 160)$, uniformly in a .

Proof. Let $A = t + L_n$ and $G = A^{-1}$, and let B be the matrix with $B_{xj} = 1$ when $x \in V$ and $j \notin V$ are neighbours. The law $\mathcal{W}(\cdot | \eta)$ has covariance $b^{-1}G$ and mean $GB\eta$. The matrix A has nonpositive off-diagonal entries and is positive definite, so $G \geq 0$ entrywise. Changing η_j by Δ translates the law by ΔGB_j , where B_j is the column of j . The coupling by translation gives $W_1^{dv} \leq |\Delta| \sum_x (GB_j)_x$, and the 1-Lipschitz function $\sum_x \phi_x$ gives the reverse inequality, so $C_{V,j} = \mathbf{1}^\top GB_j$. Since $B\mathbf{1} = L_n\mathbf{1}$ (both count the exterior neighbours of each site),

$$\sum_{j \notin V} C_{V,j} = \mathbf{1}^\top GL_n \mathbf{1} = \mathbf{1}^\top \mathbf{1} - t \mathbf{1}^\top G \mathbf{1} = n^d \alpha_n(t).$$

The reflection $x_i \mapsto n - 1 - x_i$ preserves V and the coefficients and, as n is even, exchanges the two residues of the coordinate x_i ; so the 2^d class sums are equal, and each is $n^d \alpha_n(t) / 2^d = \alpha_n(t)N_n$. This proves the formula and (i).

For (ii) let $v(i) = (i+1)(n-i)$ for $0 \leq i < n$ and $\psi(x) = \prod_i v(x_i)$. The one-dimensional Dirichlet Laplacian maps v to the constant 2, so with $s_1 = \sum_i v(i) = n(n+1)(n+2)/6$ and $s_2 = \sum_i v(i)^2 = n(n+1)(n+2)(n^2+2n+2)/30$,

$$\mathbf{1}^\top \psi = s_1^d, \quad \psi^\top A \psi = t s_2^d + 2d s_1 s_2^{d-1}.$$

By the Cauchy–Schwarz inequality for the form of A^{-1} , $\mathbf{1}^\top G \mathbf{1} \geq (\mathbf{1}^\top \psi)^2 / \psi^\top A \psi$, hence

$$1 - \alpha_n(t) \geq \left(\frac{s_1^2}{ns_2} \right)^d \frac{t}{t + 2ds_1/s_2} \geq \left(\frac{5}{6} \right)^d \frac{tn^2}{tn^2 + 10d},$$

since $s_1^2/(ns_2) = \frac{5}{6}(n^2 + 3n + 2)/(n^2 + 2n + 2) \geq \frac{5}{6}$ and $2ds_1/s_2 = 10d/(n^2 + 2n + 2) \leq 10d/n^2$. With $d = 4$ and $tn^2 = m^2(na)^2 \geq (m\ell)^2/4$ the last fraction is at least $(m\ell)^2/((m\ell)^2 + 160)$. $\square$

The Gaussian field is not Yang–Mills theory, and the proposition is a benchmark for the condition, not a statement about the Wilson measure. It shows that the averaged cube condition is met at a fixed physical scale by a field with a mass and fails exactly, at every size and every stiffness, for a massless one. Small local fluctuations alone therefore give nothing: in the massless Gaussian field on $\mathbb{Z}^d$, $d \geq 3$, the variance of

$\phi_x - \phi_{x+e_1}$ is $1/(db)$, which tends to zero as $b \rightarrow \infty$, and still $K_c = N_n$. A proof of (D) along the tuning has to produce in the gauge-fixed variables the loss that the parameter t supplies here, from the Wilson action alone.

The quadratic approximation of the forest gauge itself, with the plaquette action in place of the Laplacian, is no better than the massless field of Proposition 62(i).

Proposition 63. Let $b > 0$, let n be even and V the cube of Definition 42. For real functions ϕ on the links of $\mathbb{Z}^4$ that vanish on F , let $(d\phi)_p$ be the sum of ϕ around the plaquette p with the signs of its orientation. Given $\eta \in \mathbb{R}^{A \setminus V}$, let $\mathcal{W}(\cdot | \eta)$ be the law on $\mathbb{R}^V$ with density proportional to

$$\exp\left(-\frac{b}{2} \sum_{p \cap V \neq \emptyset} (d\phi)_p^2\right),$$

with $\phi = \eta$ on $A \setminus V$ and $\phi = 0$ on F , and define $C_{V,j}$ and K_c as above with $d_V(\phi, \phi') = \sum_{i \in V} |\phi_i - \phi'_i|$ and $\delta(\eta_j, \eta'_j) = |\eta_j - \eta'_j|$. Then

$$\sum_c K_c \geq |V| = 48N_n,$$

so $K_c \geq N_n$ for some class c , and $D(n, \alpha)$ fails for every $\alpha < 1$, at every n and every b .

Proof. The laws. The quadratic form $\sum_{p \cap V \neq \emptyset} (d\phi)_p^2$ with $\phi = 0$ outside V vanishes only at $\phi = 0$: the argument of Corollary 72, written additively, gives a function h on the sites with $\phi_{(x,\mu)} = h(x) - h(x + e_\mu)$, constant outside the box and along the lines of F , hence constant. So $\mathcal{W}(\cdot | \eta)$ is a Gaussian law whose covariance does not depend on η and whose mean is $\sum_j \eta_j u_j$ for some vectors $u_j \in \mathbb{R}^V$. Two exteriors that differ in finitely many links give translates of one law by $t = \sum_j (\eta'_j - \eta_j) u_j$; the coupling by translation and the 1-Lipschitz function $\sum_i \text{sign}(t_i) \phi_i$ show that their distance is $\|t\|_1$. In particular $C_{V,j} = \|u_j\|_1$, and

$$W_1^{dv}(\mathcal{W}(\cdot | \eta), \mathcal{W}(\cdot | \eta')) \leq \sum_j C_{V,j} |\eta'_j - \eta_j|.$$

Lines. By Lemma 9(i) the lines of F are straight, disjoint and cover the sites, and the two ends of an active link (x, μ) lie on different lines, since the line through x has direction $\sigma(x) \neq \mu$. For a line ℓ let h_ℓ be the indicator of its sites and $(dh_\ell)_{(x,\mu)} = h_\ell(x) - h_\ell(x + e_\mu)$. Then dh_ℓ vanishes on F and around every plaquette, so the density is invariant under $\phi \mapsto \phi + dh_\ell$, and $\mathcal{W}(\cdot | dh_\ell)$ is $\mathcal{W}(\cdot | 0)$ translated by the restriction of dh_ℓ to V . Hence

$$\sum_{i \in V} |(dh_\ell)_i| \leq \sum_{j \in A \setminus V} C_{V,j} |(dh_\ell)_j|.$$

Here $|(dh_\ell)_e| = 1$ when exactly one end of e lies on ℓ and 0 otherwise, so $\sum_\ell |(dh_\ell)_e| = 2$ for every active link e , the sum running over the finitely many lines that meet a link entering the inequality. Summing over these lines gives $2|V| \leq 2 \sum_j C_{V,j} = 2 \sum_c K_c$. $\square$

The proof uses only that gauge transformations constant on the lines of F move the whole cube while changing few boundary links, which is the behaviour of gauges of axial type anticipated in Remark 61. The class sums themselves, computed exactly for $n \leq 8$, lie between 1.57 and 11.24 times N_n and increase with n . Below the correlation length the theory is weakly coupled, and the quadratic approximation is its leading order there. A proof of (D) along the tuning therefore has to show that the nonperturbative regime reduces the cube coefficients of the forest gauge below N_n , or (D) has to be posed for other variables.

Remark 64. Five other routes stop at (D) or (E), or earlier.

- (a) *Disjoint blocks.* The block criterion of Lemma 114 with a partition into cubes of side n requires, for a link on a face, that its influence on the neighbouring cube be less than one. In the count of Remark 61 that influence is of order ξ , which grows like $1/a$. Averaging over the N_n translates divides it by the number of translates on whose boundary the link lies, a factor of order $1/n$, and this is why (D) is stated for averaged cubes.
- (b) *Rescaling a fixed lattice gap.* Any bound uniform in lattice units, such as those of Theorems 1–7, gives trivial continuum two-point functions [15]; rescaling it produces no gap.
- (c) *Assuming the gap.* Hypothesis R of [15] assumes the gap in physical units together with the limit. Theorem 45 replaces the gap part by the cube estimate (D), which concerns one box and one coupling at each spacing.
- (d) *Computation.* At a fixed coupling and a fixed small n , $D(n, \alpha)$ is a finite set of Kantorovich bounds and could in principle be certified as in Propositions 34 and 35; for $n = 2$ and $\text{SU}(2)$ the block already has 48 links and 144 real variables. (D) asks for it at infinitely many couplings with $n \rightarrow \infty$, which no finite computation covers.
- (e) *Axioms alone.* Proposition 41 shows that positivity, invariance and regularity do not force a gap, and Proposition 40 shows that the growth route to rotation invariance is closed for the action density. The infrared input therefore has to be an estimate such as (D).

B. Estimates uniform in the coupling

Some local quantities of the Wilson measure can be controlled uniformly in β_W , the volume and the boundary condition, by integrating by parts in a single link. Such estimates hold along every tuning, but they concern local fields in lattice units, and the three results below show how far they are from the open inputs: the force has bounded moments but its centred form is a null vector of the reflected form, the plaquette has fluctuations of order β_W^{-2} but this is a single-site variance, and bare links over a fixed physical distance become independent Haar variables.

Let $G = \text{SU}(N)$, write $b = \beta_W$, so that the Wilson weight is e^{bS} with $S = \sum_p N^{-1} \text{Re Tr } Q_p$, and give $\mathfrak{g}$ the metric $\langle T, R \rangle =$

$-\text{Tr}(TR)$. For a link e let S_e be the sum of the six plaquette terms containing e , X_T the derivative along $U_e \mapsto e^{uT} U_e$ at $u = 0$, ∇_e and Δ_e the gradient and Laplacian in U_e , $q = N^2 - 1$, $c_N = (N^2 - 1)/N$, and

$$E_e = b|\nabla_e S_e|^2, \quad W_e = E_e - c_N S_e = b|\nabla_e S_e|^2 + \Delta_e S_e.$$

Below, ω is a Wilson Gibbs state at $b > 0$: a finite-volume measure in which e is not fixed by the boundary condition, or an infinite-volume state whose conditional law of U_e given the other links is proportional to $e^{bS_e} dU_e$ with dU_e the Haar measure.

Proposition 65. (i) For every unit $T \in \mathfrak{g}$, $t \in \mathbb{R}$ and $0 \leq \eta < N/12$,

$$\begin{aligned} \mathbb{E}_\omega [e^{t\sqrt{b}X_T S_e} \mid U_{e^c}] &\leq e^{3t^2/N}, \\ \mathbb{E}_\omega [e^{\eta E_e} \mid U_{e^c}] &\leq (1 - 12\eta/N)^{-q/2}, \end{aligned}$$

so $\mathbb{E}_\omega E_e^m \leq m!(24/N)^m 2^{q/2}$ for $m \geq 1$, uniformly in b , the volume and the boundary condition. Moreover $\mathbb{E}_\omega E_e = c_N \mathbb{E}_\omega S_e$.

(ii) For a bounded function F of finitely many links, continuously differentiable in U_e , $\mathbb{E}_\omega [W_e F] = -\mathbb{E}_\omega \langle \nabla_e S_e, \nabla_e F \rangle$. In particular $\mathbb{E}_\omega [W_e F] = 0$ if F does not depend on U_e , and $\mathbb{E}_\omega [W_e W_f] = 0$ if no plaquette contains both e and f .

(iii) Let ω be reflection positive as in Appendix F, and let $A = \sum_e h_e W_e$ be a finite sum over links whose endpoints have times at least 2. Then $\omega(\vartheta(\bar{A})A) = 0$, so $\Psi_A = 0$: in the reconstructed Hilbert space the packets $\sum_e h_e E_e$ and $c_N \sum_e h_e S_e$ define the same vector.

Proof. (i) Along $u \mapsto e^{uT} U_e$ a plaquette term containing e is $N^{-1} \text{Re Tr}(e^{uT} V)$ or $N^{-1} \text{Re Tr}(V e^{-uT})$ with V unitary, whose second derivative is $N^{-1} \text{Re Tr}(T^2 e^{\pm uT} V)$. Since $-T^2 \geq 0$ and $\text{Tr}(-T^2) = 1$, $|\text{Tr}(T^2 M)| \leq 1$ for unitary M , so the second derivative of S_e along the curve is at most $6/N$ in absolute value. Taylor's formula with $u = t/\sqrt{b}$ gives $b(S_e(e^{tT/\sqrt{b}} U_e) - S_e(U_e)) \geq t\sqrt{b}X_T S_e(U_e) - 3t^2/N$. The conditional density is e^{bS_e}/Z with respect to Haar measure, and the left invariance of Haar measure gives $\mathbb{E}_\omega [e^{b(S_e(e^{tT/\sqrt{b}} U_e) - S_e(U_e))} \mid U_{e^c}] = 1$, which with the Taylor bound proves the first inequality. For the second let G_0 be a standard Gaussian vector in $\mathfrak{g}$, independent of ω ; then $e^{\eta|Y|^2} = \mathbb{E} e^{\sqrt{2\eta}\langle G_0, Y \rangle}$ for $Y = \sqrt{b}\nabla_e S_e$, and the first inequality in the direction of G_0 with $t = \sqrt{2\eta}|G_0|$ gives the bound $\mathbb{E} e^{6\eta|G_0|^2/N} = (1 - 12\eta/N)^{-q/2}$. With $\eta = N/24$ and $x^m \leq m! \eta^{-m} e^{\eta x}$ the moment bound follows. For the mean, $\sum_A T_A^2 = -c_N I$ in an orthonormal basis gives $\Delta_e S_e = -c_N S_e$, and the integral of the divergence of $e^{bS_e} \nabla_e S_e$ over G vanishes.

(ii) The integral over U_e of the divergence of $e^{bS_e} F \nabla_e S_e$ vanishes, and that divergence is $e^{bS_e} (W_e F + \langle \nabla_e S_e, \nabla_e F \rangle)$; integrate the conditional identity. The function W_f depends only on the links of the plaquettes containing f .

(iii) Each W_e in the sum depends on links with times at least 1, so $\vartheta(\bar{A})$ depends only on links with times at most 0, none of which is a link e of the sum. By (ii), $\omega(\vartheta(\bar{A})W_e) = 0$ for each e . $\square$

Proposition 66. *Let $G = \text{SU}(2)$, $b \geq 0$, ω a Wilson Gibbs state at b and p a plaquette with a link that is not fixed by the boundary condition. Then*

$$\text{Var}_\omega\left(\frac{1}{2} \text{Tr} Q_p\right) \geq \frac{1}{2^{19}(1+6b)^2}, \quad \mathbb{E}_\omega\left[1 - \frac{1}{2} \text{Tr} Q_p\right] \geq \frac{1}{4b+1}.$$

Proof. Let e be the link and identify $\text{SU}(2)$ with the unit quaternions. Given the other links, $U_e = x$ has density proportional to $e^{rx \cdot u}$ with $|u| = 1$ and $0 \leq r \leq 6b$, since the six staples have norm one, and $\frac{1}{2} \text{Tr} Q_p = x \cdot v$ for a unit vector v fixed by the conditioning. Let $X = x \cdot u$ and $T = 1 - X$; the law of T has density proportional to $e^{-rt} \sqrt{t(2-t)}$ on $[0, 2]$. Put $n = 1 + r$. Since $e^{-rt} \leq e^2 e^{-nt}$, the normalisation is at most $e^2 \int_0^\infty \sqrt{2t} e^{-nt} dt = e^2 \sqrt{\pi/2} n^{-3/2} < 14n^{-3/2}$. On $I = [\frac{1}{4n}, \frac{1}{2n}]$ the density is at least $\frac{1}{2} n^{-1/2} e^{-1/2}$ and on $J = [\frac{1}{n}, \frac{3}{2n}]$ at least $\frac{1}{2} n^{-1/2} e^{-3/2}$, so $P(T \in I) \geq \frac{1}{224}$ and $P(T \in J) \geq \frac{1}{336}$. With an independent copy T' , $\text{Var} T = \frac{1}{2} \mathbb{E}(T - T')^2 \geq P(T \in I)P(T \in J)(2n)^{-2} > 2^{-19} n^{-2}$. By symmetry the conditional covariance of x is diagonal in u and $u^\perp$, with eigenvalues $\text{Var} X$ and $\mathbb{E}(1 - X^2)/3$; on I , $1 - X^2 = T(2 - T) \geq \frac{5}{8n}$, so $\mathbb{E}(1 - X^2)/3 \geq \frac{1}{1792n} \geq 2^{-19} n^{-2}$. Hence the conditional variance of $x \cdot v$ is at least $2^{-19}(1+6b)^{-2}$, and so is the variance, by the law of total variance. For the mean, integrating $\frac{d}{dX}(e^{rX}(1 - X^2)^{3/2})$ over $[-1, 1]$ gives $3\mathbb{E}X = r\mathbb{E}(1 - X^2) \leq r(1 - (\mathbb{E}X)^2)$, whence $\mathbb{E}X \leq (\sqrt{9+4r^2} - 3)/(2r) \leq 2r/(2r+3)$ for $r > 0$, and $\mathbb{E}[1 - x \cdot v] \geq 1 - \mathbb{E}X \geq 3/(2r+3) \geq 1/(4b+1)$. $\square$

For the blocking statement let $G = \text{SU}(2)$, $B \geq 4$ an integer and, for a coarse edge (x, μ) , $V_{(x, \mu)} = \prod_{r=0}^{B-1} U_{(Bx+r\hat{\mu}, \mu)}$ the ordered product of the fine links along it.

Proposition 67. *For every Wilson Gibbs state ω of $\text{SU}(2)$ at $b > 0$ and every set $\mathcal{E}$ of K coarse edges,*

$$\begin{aligned} \|\text{Law}_\omega((V_e)_{e \in \mathcal{E}}) - m^{\otimes K}\|_{\text{TV}} &\leq \frac{K}{2} e^{3b} \left(\frac{4b}{4b+1}\right)^{B-3} \\ &\leq \frac{K}{2} \exp\left(3b - \frac{B-3}{4b+1}\right), \end{aligned}$$

where m is the Haar measure. If F is a function of the V_e of a set of coarse edges at positive times, $|F| \leq 1$, and ε is this bound for those edges and their reflections, then $0 \leq \omega(\vartheta(\bar{F}^c)F^c) \leq 6\varepsilon$ with $F^c = F - \omega(F)$.

Proof. Select on each coarse edge the fine links with $1 \leq r \leq B-2$. No plaquette contains two selected links: collinear links do not share a plaquette, parallel coarse edges are B apart, and a plaquette containing a selected link of (x, μ) has its links in direction $v \neq \mu$ at μ -offsets $r, r+1 \in [1, B-1]$ from Bx , while coarse edges in direction v lie at offsets in $B\mathbb{Z}$. Given the other links, the selected ones are therefore independent, each with density $e^{r_i x \cdot u_i} / Z(r_i)$, $0 \leq r_i \leq 6b$. Convolution with such a density acts on the representation of spin $l \geq 1$ by $D^{(l)}(u_i)$ times $I_{l+1}(r_i)/I_l(r_i)$, whose absolute value is at most $I_2(r_i)/I_1(r_i)$ because $I_\nu(r)$ decreases in $\nu \geq 0$, and $I_2(r_i)/I_1(r_i) = \mathbb{E}X \leq 4b/(4b+1)$ by the proof of Proposition 66, so it contracts the mean-zero part of L^2 by that factor. The density is at most e^{6b} because $Z \geq 1$ by Jensen's inequality, so its distance to 1 in L^2

is at most e^{3b} . The conditional law of V_e is that of a product of $B-2$ such independent variables between two fixed ones, so its total variation distance to m is at most $\frac{1}{2} e^{3b} (4b/(4b+1))^{B-3}$, and the coarse edges are conditionally independent. Total variation distance from a product is at most the sum over the factors, and averaging over the conditioning does not increase it. For the last statement, under $m^{\otimes K}$ the covariance vanishes because F and its reflection depend on disjoint edges; replacing ω by $m^{\otimes K}$ changes $\omega(\vartheta(\bar{F})F)$ by at most 2ε and $|\omega(F)|^2$ by at most 4ε . $\square$

With $B = \ell/a_k$ for a fixed physical length ℓ and $b_k = O(\log(1/a_k))$ the bound tends to zero: the bare coarse links of a fixed coarse graph become independent Haar variables, their reflected form has rank one, and a renormalisation Z_k that keeps the reflected norm of $Z_k F^c$ at least $c > 0$ must grow at least like $(c/6\varepsilon)^{1/2}$. The renormalised flowed fields of Theorem 48 are not bounded functions of bare coarse links, and none of the three propositions bears on the slice ratio, the tree bounds or (D).

C. What the cube estimate requires

This subsection proves seven facts about $D(n, \alpha)$. It holds at small coupling for every even n above an explicit size (Proposition 69), and it survives small perturbations of the action (Proposition 70), although gauge-invariant perturbations still need strong coupling at the scale of the cube (Proposition 77). It forces the gauge-fixed variables of a cube with trivial exterior to stay a fixed distance from the identity on average (Theorem 71, and Proposition 73 for $\text{SU}(2)$), so that for every fixed n it fails once β is large (Corollary 72). And it cannot be obtained from bounds on the conditional law of a single line of links that are uniform in the configuration around the line (Proposition 81). For the compact abelian theory with the Villain action it fails at every size once β is large (Proposition 78). In Sec. XII A the estimate is posed for block averages, where its quadratic approximation holds along the tuning. None of these decides (D) along the continuum tuning, where n grows while β grows.¹⁷

For a finite $V \subset A$ and $j \in A \setminus V$ let $b(V, j)$ be the number of links of V that share a plaquette with j , and write $k \sim i$ when two active links share a plaquette.

Lemma 68. *Suppose that $C_{ee'} \leq c_0$ for all active e, e' , with $12c_0 < 1$. Then $C_{V,j} \leq c_0 b(V, j)/(1 - 12c_0)$ for every finite $V \subset A$ and every $j \in A \setminus V$.*

Proof. Fix Q, Q' that differ only at j , put $d = \delta(Q_j, Q'_j)$, $\mu = \gamma_V(\cdot | Q)$ and $\nu = \gamma_V(\cdot | Q')$. The couplings of μ and ν form a weakly compact set on which $\pi \mapsto \int d\nu d\pi$ is continuous, so

¹⁷ The four statements answer the question of which parts of (D) can be settled with the methods of this paper. The first places (D) in the range where the single-link method already works, the second and third show what any proof of (D) at weak coupling has to produce, and the fourth excludes the most direct one-dimensional route.

one of them, π , minimises it; let $\rho_i = \int \delta(X_i, Y_i) d\pi$. Fix $i \in V$. Under μ the conditional law of X_i given the other links of V is $\gamma_i^F(\cdot | X)$, and likewise for v . Replace (X_i, Y_i) by a pair drawn from an optimal coupling of $\gamma_i^F(\cdot | X)$ and $\gamma_i^F(\cdot | Y)$, chosen measurably [75, Cor. 5.22], and keep the other coordinates. The new coupling still has marginals μ and v , and changing the links in which X and Y differ one at a time gives

$$W_1(\gamma_i^F(\cdot | X), \gamma_i^F(\cdot | Y)) \leq \sum_{k \in V, k \neq i} C_{ik} \delta(X_k, Y_k) + C_{ij} d.$$

Minimality of π therefore gives $\rho_i \leq c_0 \sum_{k \in V, k \sim i} \rho_k + c_0 [i \sim j] d$. Sum over $i \in V$: every ρ_k occurs at most twelve times by Lemma 9(iv), so $(1 - 12c_0) \sum_i \rho_i \leq c_0 b(V, j) d$. Since $W_1^{dv}(\mu, v) \leq \sum_i \rho_i$, dividing by d proves the lemma. $\square$

Proposition 69. *Let $n = 2m$ and let V be the cube of Definition 42. For every class c ,*

$$\sum_{j \in c \setminus V} b(V, j) \leq 10m^3 - 2m^2,$$

with equality for three classes. Consequently, if $C_{ee'} \leq c_0$ for all e, e' with $12c_0 < 1$, then $D(n, \alpha)$ holds for every $\alpha \in (0, 1)$ with

$$\alpha \geq \hat{\alpha}_n = \frac{c_0}{1 - 12c_0} \left(\frac{20}{n} - \frac{8}{n^2} \right).$$

For $SU(2)$ and $|\beta_W| < \frac{1}{3}$ this reads $\hat{\alpha}_n = \frac{|\beta_W|}{1 - 3|\beta_W|} \left(\frac{5}{n} - \frac{2}{n^2} \right)$, which is below one for every even $n > 5|\beta_W|/(1 - 3|\beta_W|)$.

Proof. Let $j = (y, \mu)$ lie in the class c , with $y \equiv r$ modulo two. A link of V that shares a plaquette with j has base point $y + d$ with $d \in \{-1, 0, 1\}^4$ and at most two nonzero coordinates, and the twelve offsets d depend only on c . For a fixed offset we count the $y \equiv r$ with $y + d \in [0, n]^4$ and $y \notin [0, n]^4$. In the coordinate l there are m values of $y_l \equiv r_l$ with $y_l + d_l \in [0, n]$, and one of them lies outside $[0, n]$ exactly when $d_l = 1$ and $r_l = 1$ (the value $y_l = -1$) or $d_l = -1$ and $r_l = 0$ (the value $y_l = n$). If q is the number of such coordinates, the count is $m^4 - m^{4-q}(m-1)^q$, that is 0, m^3 or $2m^3 - m^2$ for $q = 0, 1, 2$. Summing over the twelve offsets, the class sum is $a_c m^3 - b_c m^2$, where b_c is the number of offsets with $q = 2$ and a_c is the number with $q = 1$ plus twice b_c . Reading the offsets from Table III, the pairs (a_c, b_c) over the 48 classes are $(2, 0)$, $(7, 0)$, $(7, 1)$, $(9, 1)$, $(9, 2)$ and $(10, 2)$ three times each and $(3, 0)$, $(4, 0)$, $(5, 0)$, $(6, 0)$ and $(8, 1)$ six times each. In every case $a_c \leq 10$ and $8 - a_c + b_c \geq 0$, so

$$(10m^3 - 2m^2) - (a_c m^3 - b_c m^2) = m^2((10 - a_c)m + b_c - 2) \geq m^2(8 - a_c + b_c) \geq 0,$$

with equality for the three classes with $(a_c, b_c) = (10, 2)$. By Lemma 68, $K_c \leq c_0(10m^3 - 2m^2)/(1 - 12c_0)$, and $N_n = m^4$ gives $K_c/N_n \leq \hat{\alpha}_n$. For $SU(2)$ the proof of Theorem 1 gives $c_0 = |\beta| = |\beta_W|/4$. $\square$

At $\beta_W = \frac{3}{10}$ and $n = 16$, for instance, $\hat{\alpha}_{16} = \frac{117}{128}$, and Theorem 43 gives the gap $\log \frac{128}{117}/32$, much weaker than Theorem 1.

The proposition shows that the cube condition is not empty; it does not extend the range, because the absorption $12c_0 < 1$ is the single-link condition itself.

The cube condition is also stable under small changes of the action, which is what a block renormalisation would need at its last scale, where the effective action is no longer of Wilson form. Let $r \geq 0$ and $\varepsilon \geq 0$, and let (w_X) be real functions on G^A indexed by finite sets $X \subset A$ of diameter at most r , with w_X depending only on the links of X , such that the family is invariant under the translations by $2\mathbb{Z}^4$, every active link lies in at most M of the sets X , every X has at most s links, and

$$|w_X(Q) - w_X(Q')| \leq \varepsilon \sum_{i \in X} \delta(Q_i, Q'_i)$$

for all Q, Q' . Let $\gamma_V^W(\cdot | Q)$ be the probability measure with density proportional to $\exp \sum_{X \cap V \neq \emptyset} w_X$ with respect to $\gamma_V(\cdot | Q)$, and define $C_{V,j}^W$, K_c^W and the condition $D^W(n, \alpha)$ as in Definition 42, with γ_V^W in place of γ_V .

Proposition 70. *Let n be even and suppose that $D(n, \alpha)$ holds at β . Then $D^W(n, \alpha')$ holds with*

$$\alpha' = \alpha e^\Omega (1 + 2\Omega) + 96s\Omega, \quad \Omega = 48MDN_n \varepsilon.$$

In particular $\alpha' < 1$ for all ε below an explicit $\varepsilon_0(n, \alpha, M, s) > 0$.

Proof. The cube has $|V| = 48N_n$ links, so d_V has diameter $\Delta_V = 48N_n D$. For fixed exterior links put $u = \sum_{X \cap V \neq \emptyset} w_X$ as a function on G^V . Each link of V lies in at most M sets X , so u is $M\varepsilon$ -Lipschitz for d_V and its oscillation is at most $M\varepsilon\Delta_V = \Omega$; we normalise u so that $\sup u = 0$, which does not change the tilted measures, and put $h = e^u$, so that $e^{-\Omega} \leq h \leq 1$ and h is $M\varepsilon$ -Lipschitz.

A common tilt. Let μ, v be probability measures on G^V and $\mu_h = h\mu/\mu(h)$, $v_h = hv/v(h)$. Let f be 1-Lipschitz for d_V ; subtracting a constant we may take $|f| \leq \Delta_V$. Then fh is $(1 + \Delta_V M\varepsilon)$ -Lipschitz, $|v(fh)| \leq \Delta_V v(h)$, $|\mu(h) - v(h)| \leq M\varepsilon W_1^{dv}(\mu, v)$ and $\mu(h) \geq e^{-\Omega}$, so

$$\begin{aligned} \mu_h(f) - v_h(f) &= \frac{\mu(fh) - v(fh)}{\mu(h)} + v(fh) \frac{v(h) - \mu(h)}{\mu(h)v(h)} \\ &\leq e^\Omega (1 + 2\Omega) W_1^{dv}(\mu, v), \end{aligned}$$

since $\Delta_V M\varepsilon = \Omega$. By Kantorovich duality $W_1^{dv}(\mu_h, v_h) \leq e^\Omega (1 + 2\Omega) W_1^{dv}(\mu, v)$.

Two tilts. If u' is another weight with $|u - u'| \leq \eta$, the density of $v_{h'}$ with respect to v_h lies in $[e^{-2\eta}, e^{2\eta}]$, so the total variation distance of the two is at most $1 - e^{-2\eta} \leq 2\eta$, and $W_1^{dv}(v_h, v_{h'}) \leq 2\eta\Delta_V$.

The coefficients. Let Q, Q' differ only at $j \notin V$, with $\delta = \delta(Q_j, Q'_j)$, and let u, u' be the weights for the two exteriors. They differ only through the sets X that contain j and meet V ; if there are m_j of them, $|u - u'| \leq m_j \varepsilon \delta$. With $\mu = \gamma_V(\cdot | Q)$ and $v = \gamma_V(\cdot | Q')$, the triangle inequality and the two steps give

$$C_{V,j}^W \leq e^\Omega (1 + 2\Omega) C_{V,j} + 2\Delta_V m_j \varepsilon.$$

The pairs (X, j) with $j \in X \setminus V$ and $X \cap V \neq \emptyset$ number at most s times the number of sets meeting V , hence at most $sM|V|$. Summing over $j \in c \setminus V$,

$$\begin{aligned} K_c^W &\leq e^\Omega(1 + 2\Omega) \alpha N_n + 2\Delta_V sM|V|\varepsilon \\ &= (\alpha e^\Omega(1 + 2\Omega) + 96s\Omega)N_n. \end{aligned}$$

The family is invariant under $2\mathbb{Z}^4$, so the coefficients are, and this is $D^W(n, \alpha')$. $\square$

The admissible perturbation is of order $1/N_n = (2/n)^4$, so the proposition is useful in a box of a fixed number of lattice spacings, which is how a renormalisation argument would use it: after enough block steps that the cube of physical size ℓ has become a cube of fixed side, the effective action has to lie within ε_0 of an action that satisfies the condition there. Whether the effective actions along the tuning do so is not decided here, and that is where the problem lies: at the scale where the blocked theory is strongly coupled enough for a condition of this kind, the running coupling is of order one, which is outside the range of the existing ultraviolet analysis. The proposition shows only that the condition itself is not the obstacle.

The next theorem uses the configuration in which every link outside V is the identity. Write $\gamma_V^1 = \gamma_V(\cdot | 1)$ and

$$\varepsilon_V(\beta) = \frac{1}{|V|} \sum_{i \in V} \int \delta(1, Q_i) d\gamma_V^1(Q)$$

for the mean distance of the gauge-fixed variables of the cube from the identity. Recall that δ is bi-invariant of diameter D .

Theorem 71. *Let n be even, V the cube of Definition 42, and suppose that $D(n, \alpha)$ holds at β . Then*

$$\varepsilon_V(\beta) \geq \frac{(1 - \alpha)D}{3}.$$

Proof. A flat configuration. By Lemma 9(i) and the period two of σ , the sixteen residues modulo two fall into eight pairs $\{r, r + e_{\sigma(r)}\}$, the residues met by one line of F . The eight residues 0010, 0011, 0100, 0110, 1001, 1011, 1100, 1101 form four of these pairs. Choose $z, w \in G$ with $\delta(z, w) = D$, put $g(x) = w$ if x is congruent to one of those eight residues and $g(x) = z$ otherwise, and let $Q_{(x, \mu)}^g = g(x)g(x + e_\mu)^{-1}$ for every link of $\mathbb{Z}^4$. Then g is constant on every line of F , so Q^g equals 1 on F , every plaquette variable of Q^g is 1, and Q^g has period two. By bi-invariance $\delta(1, Q_{(x, \mu)}^g) = \delta(g(x + e_\mu), g(x))$, which is D when x and $x + e_\mu$ receive different values and 0 otherwise; from Table III the first case occurs for exactly 32 of the 48 classes. Let δ_c be the common value of $\delta(1, Q_e^g)$ for e in the class c , so that $\sum_c \delta_c = 32D$.

Gauge covariance. Let $\Phi(Q)_{(x, \mu)} = g(x)Q_{(x, \mu)}g(x + e_\mu)^{-1}$ on G^V . The gauge transformation by g maps the configuration equal to Q_V on V and to 1 elsewhere to the one equal to $\Phi(Q_V)$ on V and to Q^g elsewhere; it leaves every plaquette trace unchanged and every link of F equal to 1, and Φ preserves Haar measure. Hence $\gamma_V(\cdot | Q^g) = \Phi_* \gamma_V^1$.

Upper bound. Only the finitely many links outside V that share a plaquette with V enter γ_V . Changing them one at a time from 1 to Q^g and applying the definition of $C_{V, j}$ at each step,

$$\begin{aligned} W_1^{d_V}(\gamma_V^1, \gamma_V(\cdot | Q^g)) &\leq \sum_{j \in A \setminus V} C_{V, j} \delta(1, Q_j^g) = \sum_c \delta_c K_c \\ &\leq \alpha N_n \sum_c \delta_c = 32\alpha D N_n. \end{aligned}$$

Lower bound. The function $f(Q) = -\sum_{i \in V} \delta(Q_i, Q_i^g)$ is 1-Lipschitz for d_V . By gauge covariance and bi-invariance, $\int f d\gamma_V(\cdot | Q^g) = -\int \sum_i \delta(Q_i, 1) d\gamma_V^1 = -|V|\varepsilon_V$, and by the triangle inequality $\int f d\gamma_V^1 \leq -\sum_i \delta(1, Q_i^g) + |V|\varepsilon_V$. Every class has N_n links in V , so $\sum_i \delta(1, Q_i^g) = 32D N_n$ and

$$W_1^{d_V}(\gamma_V^1, \gamma_V(\cdot | Q^g)) \geq 32D N_n - 2|V|\varepsilon_V.$$

With $|V| = 48N_n$ the two bounds give $96N_n\varepsilon_V \geq 32(1 - \alpha)D N_n$. $\square$

Corollary 72. *For every even n and every $\alpha < 1$ there is $\beta_0(n, \alpha)$ such that $D(n, \alpha)$ fails at every $\beta \geq \beta_0(n, \alpha)$. In particular, if $D(n_k, \alpha)$ holds at couplings $\beta_k \rightarrow \infty$, then $n_k \rightarrow \infty$.*

Proof. For $\beta > 0$, γ_V^1 has density proportional to $e^{-N\beta S}$ with respect to Haar measure, where $S = \sum_{p \cap V \neq \emptyset} (N - \text{Re Tr } Q_p^F)$ is continuous and nonnegative on G^V and vanishes exactly when $Q_p^F = 1$ for every plaquette p that meets V . We show that this forces $Q_V = 1$. The other plaquettes have all their links equal to 1, so every plaquette variable is 1; since the plaquettes of $\mathbb{Z}^4$ fill every closed lattice path, there is a function h on the sites with $Q_{(x, \mu)} = h(x)h(x + e_\mu)^{-1}$ on every link, and h takes equal values at the two ends of every link on which $Q = 1$. Every link with base point outside $[0, n]^4$ is such a link. From a site with a coordinate at least n one can move in every positive direction along links of this kind, and from a site with a negative coordinate x_V one can move in the positive directions other than v until another coordinate reaches n ; so h is constant on the sites outside the box. Every line of F leaves the box, and h is constant along it, so h is constant and $Q_V = 1$.

Let U be a neighbourhood of 1 in G^V , $s = \min_{G^V \setminus U} S > 0$, and $U' \subset U$ a neighbourhood of 1 on which $S < s/2$. Then $\gamma_V^1(G^V \setminus U) \leq e^{-N\beta s} / (|U'|e^{-N\beta s/2})$, which tends to zero, so $\varepsilon_V(\beta) \rightarrow 0$ as $\beta \rightarrow \infty$, and Theorem 71 excludes $D(n, \alpha)$ once $\varepsilon_V(\beta) < (1 - \alpha)D/3$. $\square$

For $\text{SU}(2)$ the constant of Theorem 71 can be improved: a flat configuration with four values instead of two moves every class of active links, and the means of the gauge-fixed links are scalar.

Proposition 73. *Let $G = \text{SU}(2)$ with the chordal metric, let n be even, V the cube of Definition 42, and suppose that $D(n, \alpha)$ holds at β . Then*

$$\frac{1}{|V|} \sum_{i \in V} \left| \int \text{Re Tr } Q_i d\gamma_V^1 \right| \leq 2\alpha,$$

$$\frac{1}{|V|} \sum_{i \in V} \int \delta(1, Q_i)^2 d\gamma_V^1 \geq 2(1 - \alpha),$$

and $\varepsilon_V(\beta) \geq 1 - \alpha$.

Proof. Write $\langle \cdot, \cdot \rangle$ for the Euclidean product of quaternions; then $\text{Re Tr } q = 2\langle q, 1 \rangle$ and $\delta(p, q) = |p - q|$.

Four colours. Split the sixteen residues modulo two into the sets

$$\begin{aligned} X_1 &= \{0000, 1000, 0111, 1111\}, \\ X_2 &= \{0001, 0101, 1010, 1110\}, \\ X_3 &= \{0010, 0011, 1100, 1101\}, \\ X_4 &= \{0100, 0110, 1001, 1011\}. \end{aligned}$$

Each is the union of two of the eight pairs $\{r, r + e_{\sigma(r)}\}$, and Table III shows that r and $r + e_\mu$ lie in different sets for every residue r and every $\mu \neq \sigma(r)$. (The eight residues used in the proof of Theorem 71 form $X_3 \cup X_4$.) Let

$$\begin{aligned} z_1 &= (\mathbf{i} + \mathbf{j} + \mathbf{k})/\sqrt{3}, & z_2 &= (\mathbf{i} - \mathbf{j} - \mathbf{k})/\sqrt{3}, \\ z_3 &= (-\mathbf{i} + \mathbf{j} - \mathbf{k})/\sqrt{3}, & z_4 &= (-\mathbf{i} - \mathbf{j} + \mathbf{k})/\sqrt{3}, \end{aligned}$$

unit quaternions with $\langle z_k, z_l \rangle = -\frac{1}{3}$ for $k \neq l$, so that $\delta(z_k, z_l) = d_* = (8/3)^{1/2}$. Put $g(x) = z_k$ when the residue of x lies in X_k , and $Q_{(x,\mu)}^g = g(x)g(x + e_\mu)^{-1}$. As in the proof of Theorem 71, Q^g equals 1 on F , every plaquette variable of Q^g is 1, and $\mathcal{W}(\cdot | Q^g) = \Phi_* \mathcal{W}_V^1$ with $\Phi(Q)_{(x,\mu)} = g(x)Q_{(x,\mu)}g(x + e_\mu)^{-1}$; now $\delta(1, Q_e^g) = d_*$ for every active link e .

Upper bound. Changing the exterior links one at a time,

$$W_1^{dv}(\mathcal{W}_V^1, \mathcal{W}(\cdot | Q^g)) \leq d_* \sum_c K_c \leq 48 \alpha d_* N_n = \alpha d_* |V|.$$

Lower bound. Conjugating every link of V by one unit quaternion leaves the exterior 1, the links of F , Haar measure and every plaquette trace unchanged, so it preserves $\mathcal{W}_V^1$. The mean $m_i = \int Q_i d\mathcal{W}_V^1$, taken in the space of quaternions, therefore commutes with every unit quaternion, and $m_i = a_i 1$ with $a_i = \frac{1}{2} \int \text{Re Tr } Q_i d\mathcal{W}_V^1$ real. Since Φ acts on each link by a fixed left and right multiplication, the mean of Q_i under $\mathcal{W}(\cdot | Q^g)$ is $a_i Q_i^g$. Choose unit quaternions w_i with $\langle w_i, a_i Q_i^g - a_i 1 \rangle = |a_i| |Q_i^g - 1|$. The function $f(Q) = \sum_{i \in V} \langle w_i, Q_i \rangle$ is 1-Lipschitz for dv , and Kantorovich duality gives

$$W_1^{dv}(\mathcal{W}_V^1, \mathcal{W}(\cdot | Q^g)) \geq \sum_{i \in V} |a_i| \delta(1, Q_i^g) = d_* \sum_{i \in V} |a_i|.$$

Conclusion. Comparing the two bounds, $\sum_i |a_i| \leq \alpha |V|$, which is the first inequality. Since $\delta(1, q)^2 = 2 - 2\langle q, 1 \rangle$, $\int \delta(1, Q_i)^2 d\mathcal{W}_V^1 = 2(1 - a_i) \geq 2(1 - |a_i|)$, and averaging gives the second. As $\delta \leq 2$, $\delta \geq \delta^2/2$, and the bound on ε_V follows. $\square$

The four colours need nothing from $\text{SU}(2)$ except four elements whose pairwise quotients are conjugate, and every compact simple group contains such elements. For a metric that comes from a representation the bound then compares the cube with Haar measure.

Proposition 74. *Let G be a compact simple group, ρ a non-trivial irreducible unitary representation of G of dimension d , and $\delta(p, q) = \kappa^{-1/2} \|\rho(p) - \rho(q)\|_{\text{HS}}$ for some $\kappa > 0$. Let n*

be even, V the cube of Definition 42, and suppose that $D(n, \alpha)$ holds at β . Then

$$\frac{1}{|V|} \sum_{i \in V} \left| \int \text{Tr } \rho(Q_i) d\mathcal{W}_V^1 \right| \leq \alpha d,$$

$$\frac{1}{|V|} \sum_{i \in V} \int \delta(1, Q_i)^2 d\mathcal{W}_V^1 \geq (1 - \alpha) \int_G \delta(1, q)^2 dq,$$

where dq is Haar measure, and $\varepsilon_V(\beta) \geq (1 - \alpha)D/2$.

Proof. Four elements. Let $\varphi: \text{SU}(2) \rightarrow G$ be the homomorphism of the $\text{SU}(2)$ subgroup of a root, and $z_1, \dots, z_4$ the tetrahedral quaternions of Proposition 73. For $k \neq l$ the quaternion $z_k z_l^{-1}$ has real part $\langle z_k, z_l \rangle = -\frac{1}{3}$; unit quaternions with equal real parts are conjugate in $\text{SU}(2)$, so the twelve elements $\varphi(z_k z_l^{-1})$ are conjugate in G , and by bi-invariance $\delta(\varphi(z_k), \varphi(z_l)) = d_*$ for all $k \neq l$, with one number d_* . If $d_* = 0$, then $\varphi(z_k z_l^{-1})$ lies in $\ker \rho$, a proper normal subgroup of the simple group G and therefore central; it then commutes with $\varphi(\text{SU}(2))$, so it is central there, and $z_k z_l^{-1}$ lies in $\{\pm 1\} \ker \varphi \subset \{\pm 1\}$, which its real part $-\frac{1}{3}$ excludes. So $d_* > 0$.

The two bounds. Put $g(x) = \varphi(z_k)$ when the residue of x lies in X_k . The proof of Proposition 73 applies word for word: Q^g equals 1 on F and on every plaquette, $\delta(1, Q_e^g) = d_*$ on every active link, and $W_1^{dv}(\mathcal{W}_V^1, \mathcal{W}(\cdot | Q^g)) \leq \alpha d_* |V|$. Simultaneous conjugation of the links of V preserves $\mathcal{W}_V^1$, so the mean $\int \rho(Q_i) d\mathcal{W}_V^1$ commutes with $\rho(G)$, and by Schur's lemma it is $c_i I$ with $c_i = d^{-1} \int \text{Tr } \rho(Q_i) d\mathcal{W}_V^1$. Its mean under $\mathcal{W}(\cdot | Q^g)$ is $c_i \rho(Q_i^g)$. With matrices W_i of Hilbert–Schmidt norm $\kappa^{-1/2}$ chosen so that $\text{Re Tr}(W_i^* c_i (\rho(Q_i^g) - I)) = |c_i| \delta(1, Q_i^g)$, the function $f(Q) = \sum_i \text{Re Tr}(W_i^* \rho(Q_i))$ is 1-Lipschitz for dv , and Kantorovich duality gives $W_1^{dv} \geq d_* \sum_i |c_i|$. Hence $\sum_i |c_i| \leq \alpha |V|$, the first inequality.

Comparison with Haar measure. Since $\kappa \delta(1, q)^2 = 2d - 2\text{Re Tr } \rho(q)$ and $\int_G \text{Tr } \rho(q) dq = 0$ for a nontrivial irreducible ρ , $\int_G \delta(1, q)^2 dq = 2d/\kappa$, while $\int \delta(1, Q_i)^2 d\mathcal{W}_V^1 = 2d(1 - \text{Re } c_i)/\kappa \geq 2d(1 - |c_i|)/\kappa$. Averaging gives the second inequality. Finally $\kappa D^2 \leq 4d$, because $\|\rho(p) - \rho(q)\|_{\text{HS}} \leq 2\sqrt{d}$, so $\delta \geq \delta^2/D$ gives $\varepsilon_V \geq 2d(1 - \alpha)/(\kappa D) \geq (1 - \alpha)D/2$. $\square$

For $\text{SU}(N)$ with the Hilbert–Schmidt distance this improves the constant $\frac{1}{3}$ of Theorem 71 to $\frac{1}{2}$, and for $\text{SU}(3)$, where $D^2 = 9$ and the Haar mean of δ^2 is 6, to $\frac{2}{3}$: in a cube with the identity outside, $D(n, \alpha)$ leaves the gauge-fixed links at least the fraction $1 - \alpha$ of the disorder of independent Haar-distributed links.

The size $\beta_0(n, \alpha)$ in Corollary 72 comes from a compactness argument. For $\text{SU}(2)$ it can be made explicit, and then reads as a lower bound on the side of any cube that satisfies the condition.

Proposition 75. *Let $G = \text{SU}(2)$ with the chordal metric, so that $D = 2$, let $b = \beta_W \geq \frac{1}{9}$, let $n \geq 2$ be even and put $L(b) = 1 + \log 24 + \frac{3}{2} \log(9b)$. Then*

$$\frac{1}{|V|} \sum_{e \in V} \int \delta(1, Q_e)^2 d\mathcal{W}_V^1 \leq \frac{n^2(n+2)(2n+1)L(b)}{b},$$

so $\varepsilon_V(\beta)^2$ is at most the same quantity. Consequently, if $D(n, \alpha)$ holds at b , then

$$n^4 \geq \frac{2(1-\alpha)b}{5L(b)}.$$

Proof. Put $m = n/2$. Under γ_V^1 every link outside V equals 1: the links of F always, the others by the choice of exterior. A plaquette that does not meet V therefore has plaquette variable 1.

Links as loops. Let $e = (x, \mu) \in V$, $y = x + e_\mu$, $\sigma = \sigma(x)$ and $\tau = \sigma(y)$. The directions σ, τ, μ are distinct: $\sigma \neq \mu$ because e is active; $\tau \neq \mu$ because otherwise the line of F through y would contain x , and the lines of F are disjoint; and $\tau \neq \sigma$ by Table III, in which $\sigma(x + e_\mu) \neq \sigma(x)$ whenever $\mu \neq \sigma(x)$. Since $0 \leq x_\sigma < n$, one of $n - x_\sigma$ and $x_\sigma + 1$ is at most m ; call it s and let $\varepsilon = \pm 1$ be the sign for which $x' = x + s\varepsilon e_\sigma$ has σ -coordinate outside $[0, n)$. Choose $t \leq m$ and ε' in the same way for y and τ (note $y_\tau = x_\tau$). Let R_1 be the s plaquettes of the rectangle with corners $x, x', x' + e_\mu, y$ in the plane $\{\sigma, \mu\}$, and R_2 the t plaquettes of the rectangle with corners $y, x' + e_\mu, x' + e_\mu + t\varepsilon' e_\tau, y + t\varepsilon' e_\tau$ in the plane $\{\sigma, \tau\}$. The union $R(e) = R_1 \cup R_2$ is a disc, and its boundary runs from x along the line of F through x to x' , then to $x' + e_\mu$, along e_τ to $x' + e_\mu + t\varepsilon' e_\tau$, along e_σ to $y + t\varepsilon' e_\tau$, back along the line of F through y to y , and over e to x . The links of the first and fifth pieces lie in F . The links of the second and third pieces have base points with σ -coordinate $x'_\sigma \notin [0, n)$, and those of the fourth have base points with τ -coordinate outside $[0, n)$; none of them lies in V . So the holonomy of the boundary of $R(e)$ is $Q_e^{\pm 1}$. Adding the plaquettes of a disc one at a time along its boundary shows that this holonomy is a product of conjugates of the plaquette variables $(Q_p^F)^{\pm 1}$, $p \in R(e)$, one for each plaquette. By bi-invariance and the triangle inequality,

$$\delta(1, Q_e) \leq \sum_{p \in R(e)} \delta(1, Q_p^F), \quad |R(e)| = s(t+1) \leq m(m+1).$$

Multiplicity. If a plaquette p lies in $R_2(e)$, its plane is $\{\sigma, \tau\}$, μ is one of the two other directions, and the base point of p is $y + ae_\sigma + ce_\tau$ with $-m \leq a, c \leq m-1$. Given p , there are two ways to name the directions of its plane, two choices of μ and $(2m)^2$ of (a, c) , and they determine y , hence e ; so p lies in at most $16m^2$ of the sets $R_2(e)$. Likewise $p \in R_1(e)$ forces its plane to be $\{\sigma, \mu\}$ and its base point to be $x + ae_\sigma$ with $-m \leq a \leq m-1$, which leaves at most $4m$ links e . Every plaquette therefore lies in at most $16m^2 + 4m = 4n^2 + 2n$ of the discs $R(e)$.

The energy. For the chordal metric $\delta(1, q)^2 = 2 - \text{Re Tr } q$. Let $S = \sum_{p \cap V \neq \emptyset} \delta(1, Q_p^F)^2$; then γ_V^1 has density $e^{-bS/2}/Z$ with $Z = \int e^{-bS/2}$, the integral over $\text{SU}(2)^V$ with Haar measure. The function $t \mapsto \log \int e^{-tS}$ is convex and vanishes at $t = 0$, so its derivative at $t = b/2$, which is $-\int S d\gamma_V^1$, is at least $2 \log Z/b$. A plaquette that meets V contains at most three links of V and the other links are 1, so $\delta(1, Q_p^F)^2 \leq 3 \sum_{e \in p \cap V} \delta(1, Q_e)^2$; every link lies in six plaquettes, so $S \leq 18 \sum_{e \in V} \delta(1, Q_e)^2$ and $Z \geq I^{|V|}$ with $I = \int e^{-9b \delta(1, q)^2} dq$. The chordal distance is at most the geodesic one, and the geodesic ball of radius

$r = (9b)^{-1/2} \leq 1$ has measure greater than $r^3/24$ (proof of Proposition 81), so $I \geq e^{-1}(9b)^{-3/2}/24$ and

$$\int S d\gamma_V^1 \leq \frac{2|V|L(b)}{b}.$$

Conclusion. The Cauchy–Schwarz inequality inside each disc squares the loop bound:

$$\begin{aligned} \delta(1, Q_e)^2 &\leq |R(e)| \sum_{p \in R(e)} \delta(1, Q_p^F)^2 \\ &\leq m(m+1) \sum_{p \in R(e)} \delta(1, Q_p^F)^2. \end{aligned}$$

Summing over $e \in V$ and using the multiplicity,

$$\sum_{e \in V} \delta(1, Q_e)^2 \leq m(m+1)(4n^2 + 2n)S = \frac{n^2(n+2)(2n+1)}{2} S,$$

and integrating against γ_V^1 gives the first bound; the bound on ε_V follows by the Cauchy–Schwarz inequality. If $D(n, \alpha)$ holds, Proposition 73 gives $2(1-\alpha) \leq n^2(n+2)(2n+1)L(b)/b$, and $(n+2)(2n+1) \leq 5n^2$ for $n \geq 2$. $\square$

The loops and their multiplicity use nothing of $\text{SU}(2)$ but the bi-invariance of δ , and the energy uses only the form of the action and the measure of small balls. The bound therefore holds for every $\text{SU}(N)$, with the same power of n .

Proposition 76. *Let $G = \text{SU}(N)$, $N \geq 2$, with the weight $\exp(\frac{b}{N} \sum_p \text{Re Tr } Q_p)$, $b = \beta_W \geq N/9$, and the Hilbert–Schmidt distance $\delta(p, q) = \|p - q\|_{\text{HS}}$. Let $n \geq 2$ be even and put*

$$L_N(b) = 1 + \frac{N^2 - 1}{2} \log(81\pi^2 b) + N(N-1) \log \frac{12}{11}.$$

Then

$$\frac{1}{|V|} \sum_{e \in V} \int \delta(1, Q_e)^2 d\gamma_V^1 \leq \frac{Nn^2(n+2)(2n+1)L_N(b)}{b}.$$

Consequently, if $D(n, \alpha)$ holds at b , then

$$n^4 \geq \frac{2(1-\alpha)b}{5L_N(b)}.$$

Proof. Loops. The discs $R(e)$ of the proof of Proposition 75 and the bound $4n^2 + 2n$ on their multiplicity do not depend on the group, and the loop bound uses only bi-invariance and the triangle inequality. With $S = \sum_{p \cap V \neq \emptyset} \delta(1, Q_p^F)^2$ the same Cauchy–Schwarz step gives $\sum_{e \in V} \delta(1, Q_e)^2 \leq \frac{1}{2} n^2(n+2)(2n+1)S$.

The energy. Since $\delta(1, q)^2 = 2N - 2\text{Re Tr } q$, the measure γ_V^1 has density $e^{-bS/(2N)}/Z$ with respect to Haar measure. As in the proof of Proposition 75, the convexity of $t \mapsto \log \int e^{-tS}$ gives $\int S d\gamma_V^1 \leq 2N \log(1/Z)/b$, and $S \leq 18 \sum_{e \in V} \delta(1, Q_e)^2$ gives $Z \geq I^{|V|}$ with $I = \int e^{-9b \delta(1, q)^2/N} dq$.

Small balls. Give $\text{su}(N)$ the inner product $\text{Tr}(X^*Y)$ and $\text{SU}(N)$ the corresponding bi-invariant metric, of volume v_N . In

exponential coordinates the Riemannian measure is $j(X)dX$, where dX is Lebesgue measure and

$$j(X) = \left| \det \frac{1 - e^{-\text{ad}X}}{\text{ad}X} \right| = \prod_{k < l} \left(\frac{\sin(\varphi_{kl}/2)}{\varphi_{kl}/2} \right)^2$$

if X has eigenvalues $i\theta_1, \dots, i\theta_N$ and $\varphi_{kl} = \theta_k - \theta_l$ [42, Ch. II, Thm. 1.7]; in particular $j \leq 1$. Every $q \in \text{SU}(N)$ is conjugate to a diagonal matrix with entries $e^{i\psi_k}$, $-\pi \leq \psi_k < \pi$; then $\sum_k \psi_k = 2\pi k_0$ with $|k_0| \leq N/2$, and changing $|k_0|$ of the ψ_k by $\mp 2\pi$ gives angles θ_k with sum zero and $|\theta_k| \leq 3\pi$. So $q = e^X$ with $X \in \mathfrak{su}(N)$ and $\|X\|_{\text{HS}} \leq 3\pi\sqrt{N}$, and $v_N \leq \omega_{N^2-1} (3\pi\sqrt{N})^{N^2-1}$, where ω_k is the volume of the unit ball of $\mathbb{R}^k$. Now let $\|X\|_{\text{HS}} \leq r \leq 1$. The eigenvalues have $|\theta_k| \leq 1 < \pi$, so $\exp$ is injective on this ball; $\delta(1, e^X)^2 = \sum_k |e^{i\theta_k} - 1|^2 \leq \|X\|_{\text{HS}}^2$; and $|\varphi_{kl}|/2 \leq (\theta_k^2 + \theta_l^2)^{1/2}/\sqrt{2} \leq 2^{-1/2}$, so each factor of $j(X)$ is at least $(1 - \frac{1}{2})^2$ because $\sin x/x \geq 1 - x^2/6$. Dividing by v_N , the Haar measure of $\{q : \delta(1, q) \leq r\}$ is at least

$$\left(\frac{11}{12}\right)^{N(N-1)} \frac{\omega_{N^2-1} r^{N^2-1}}{v_N} \geq \left(\frac{11}{12}\right)^{N(N-1)} \left(\frac{r}{3\pi\sqrt{N}}\right)^{N^2-1}.$$

With $r = (N/9b)^{1/2} \leq 1$ this gives $I \geq e^{-1}$ times the right-hand side, and since $3\pi\sqrt{N}/r = 9\pi\sqrt{b}$, $\log(1/I) \leq L_N(b)$. Hence $\int S d\gamma_V^1 \leq 2N|V|L_N(b)/b$, and the first bound follows from the loops.

Conclusion. Proposition 74 with the defining representation and $\kappa = 1$, for which the Haar mean of $\delta(1, q)^2$ is $2N$, gives $2N(1-\alpha) \leq Nn^2(n+2)(2n+1)L_N(b)/b$ when $D(n, \alpha)$ holds, and $(n+2)(2n+1) \leq 5n^2$. $\square$

For every N , a cube on which the condition holds therefore has side at least of order $(\beta_W / \log \beta_W)^{1/4}$, with a constant that depends on N only through L_N , which grows like $\frac{N^2}{2} \log \beta_W$.

The cube condition is a statement about strong coupling at its own scale, and this persists under the perturbations of Proposition 70 when they respect the gauge. Call (w_X) gauge invariant if $w_X(Q^h) = w_X(Q)$ for every function h on the sites that is constant on every line of F , where $Q_{(x,\mu)}^h = h(x)Q_{(x,\mu)}h(x+e_\mu)^{-1}$. Such transformations keep the links of F equal to 1, so they act on G^A , and the restriction to the gauge of an action invariant under all gauge transformations is gauge invariant in this sense. Write ε_V^W for ε_V with $\gamma_V^W(\cdot | 1)$ in place of γ_V^1 .

Proposition 77. *Let (w_X) be as in Proposition 70 and gauge invariant, let $n = 2m$, and suppose that $D^W(n, \alpha)$ holds.*

$$(i) \ \varepsilon_V^W \geq (1-\alpha)D/3.$$

(ii) *If $\eta \geq 0$, $0 \leq \kappa \leq 1$, and the event that $\delta(1, Q_p^F) \leq \eta$ for every plaquette p that meets V has probability at least $1 - \kappa$ under $\gamma_V^W(\cdot | 1)$, then*

$$m(m+1)\eta + \kappa D \geq \frac{(1-\alpha)D}{3}.$$

Proof. (i) The proof of Theorem 71 uses the action only through the identity $\gamma_V(\cdot | Q^g) = \Phi_* \gamma_V^1$ and through the coefficients $C_{V,j}$. The function g of that proof is constant on

every line of F , so the gauge transformation by g leaves every w_X unchanged, as it leaves every plaquette trace, and the same identity holds for γ_V^W . Only finitely many links outside V enter γ_V^W , because the sets X have bounded diameter, so the upper bound holds with $C_{V,j}^W$ and K_c^W in place of $C_{V,j}$ and K_c ; the lower bound is unchanged.

(ii) Under $\gamma_V^W(\cdot | 1)$ every link outside V equals 1, so a plaquette that does not meet V has variable 1. The loop bound of the proof of Proposition 75, which uses only the bi-invariance of δ , gives $\delta(1, Q_e) \leq \sum_{p \in R(e)} \delta(1, Q_p^F) \leq m(m+1)\eta$ for every $e \in V$ on the event of the statement, and $\delta(1, Q_e) \leq D$ off it. Averaging over $e \in V$ and integrating, $\varepsilon_V^W \leq m(m+1)\eta + \kappa D$, and (i) completes the proof. $\square$

In the small-field region of a gauge-covariant block renormalisation the blocked plaquette variables lie within a distance η that tends to zero with the running coupling, and large fields have small probability κ . If $\kappa < (1-\alpha)/3$, part (ii) allows the cube condition on the blocked lattice only when $m(m+1) \geq x$, where $x = ((1-\alpha)/3 - \kappa)D/\eta$, so the side must grow like $\eta^{-1/2}$. On such a cube $N_n = m^4 \geq x^2/4$, and since $\alpha' \geq \alpha + 96s\Omega$, Proposition 70 gives $\alpha' < 1$ only for

$$\begin{aligned} \varepsilon &< \frac{1-\alpha}{4608sMDN_n} \\ &\leq \frac{(1-\alpha)\eta^2}{1152sMD^3((1-\alpha)/3 - \kappa)^2}, \end{aligned}$$

which is of order η^2 . At a fixed side the condition is therefore a statement about strong coupling, and a renormalisation argument has to carry the blocked theory out of the small-field regime before the condition can be checked. That passage, at a running coupling of order one, is the step that remains open.

The next statement shows that the cube condition separates a theory with a massless photon from the expected behaviour of $\text{SU}(N)$: for the compact abelian theory in its Coulomb phase it fails at every scale. Let $G = \text{U}(1)$ with the chordal metric, $D = 2$, and replace the Wilson weight by the Villain weight

$$\varphi_\beta(\theta) = \sum_{k \in \mathbb{Z}} e^{-\beta(\theta + 2\pi k)^2/2},$$

so that the measure of a box Λ has density proportional to $\prod_{p \in P(\Lambda)} \varphi_\beta(\theta_p)$, where $e^{i\theta_p}$ is the plaquette variable. Define γ^F , $C_{V,j}$, K_c and $D(n, \alpha)$ as before with φ_β in place of the Wilson weight.

Proposition 78. *There is β_0 such that for the Villain theory at every $\beta > \beta_0$ the condition $D(n, \alpha)$ fails for every even $n \geq 2$ and every $\alpha < 1$. In particular, for couplings $\beta_k \rightarrow \infty$ and any even n_k , the hypothesis of Theorem 45 fails for this theory.*

Proof. Lemma 10, Proposition 11(i)–(iii) and (v), Lemma 112, Corollary 113 and Lemma 115 use the Wilson weight only as a positive continuous class function of the plaquette variable, with one exception: the Doeblin bound in Lemma 115(a), which for the Villain weight reads that each single-link density is at least $(\min \varphi_\beta / \max \varphi_\beta)^6 > 0$ times its maximum. Suppose that $D(n, \alpha)$ holds. As in the proof of Theorem 43, the

limit $\omega(O) = \lim W_\Lambda(O)$ of the measures of boxes with free boundary conditions then exists along every sequence of boxes exhausting $\mathbb{Z}^4$, and for gauge-invariant local O, O' with O Lipschitz in each link,

$$|\text{Cov}_\omega(O, O')| \leq \|O'\|_{\infty L_O} \frac{D(1 + c_n e^{\lambda \rho_n})}{1 - \theta} |S_{O'}| e^{-\lambda \text{dist}(O, O')}$$

for every $\lambda > 0$ with $\theta = \alpha e^{\lambda \rho_n} < 1$, where L_O is a Lipschitz constant of O and $S_{O'}$ the support of O' .

Put $\Psi_p = \varphi'_\beta(\theta_p)/\varphi_\beta(\theta_p)$; up to the factor $-\mathbf{i}$ this is the lattice field strength of Fröhlich and Spencer [36, Sect. 2] for the Villain action. It is real, odd and smooth in θ_p , so it is gauge invariant, local and Lipschitz in each of its four links, with constants that do not depend on p , and $\omega(\Psi_p) = 0$ by the symmetry $\theta \mapsto -\theta$. The bound above gives $\sum_{p'} |\omega(\Psi_p \Psi_{p'})| < \infty$, the sum over the plaquettes p' parallel to p . The state of [36] is the limit of the same free-boundary measures over boxes, which is ω . Fröhlich and Spencer prove, for β larger than some β_0 , that the Fourier transform of the two-plaquette function of the field strength is discontinuous at zero momentum, so that this sum is infinite [36, Sect. 2]; see also [41]. This contradiction proves the first statement, and the second follows because $\beta_k > \beta_0$ for all large k . $\square$

At large β the plaquette variables of the abelian theory are Gaussian to leading order, as are those of $\text{SU}(N)$, and every property of the action used in this paper (compactness, locality, positivity and smoothness of the weight, reflection positivity, the forest gauge) is shared by the two theories. A proof of (D) along the continuum tuning therefore has to use the non-abelian structure of G ; Proposition 78 holds at every cube size, so letting the side grow with β , which Corollary 72 already requires, does not remove the need. Fröhlich and Spencer indicate how their result extends to the Wilson action without giving the details [36, Sect. 2], so the proposition is stated for the Villain action only.

Remark 79. The proof loses the area of a disc, of order n^2 , in passing from plaquettes to links, and on loops that loss is of the right order. On the periodic lattice $(\mathbb{Z}/L\mathbb{Z})^4$ with L even and coordinates $0 \leq x_j < L$, let $\phi = 4\pi/L^2$, $q(\theta) = \cos \theta + \mathbf{i} \sin \theta$, and put $U_{(x,2)} = q(\phi x_1)$, $U_{(x,1)} = q(-\phi L x_2)$ when $x_1 = L - 1$, and $U = 1$ on every other link. Every plaquette of the plane $\{1, 2\}$ equals $q(\phi)$, including those on the seam $x_1 = L - 1$ and at the corner $x_1 = x_2 = L - 1$, where the angle is $\phi - 4\pi$; all other plaquettes equal 1. So every plaquette variable lies within ϕ of the identity, while the boundary of the square $[0, L/2]^2$ of the plane $\{1, 2\}$ has holonomy $q(\phi L^2/4) = -1$, at the largest possible distance $D = 2$, against the bound $\phi L^2/4 = \pi$ given by the disc. Every staple sum of this configuration is parallel to its link, so it is a stationary point of the Wilson action, and gradient flow [55] does not change it. Plaquette variables within $O(n^{-2})$ of the identity therefore do not force loops of length n near the identity, and an argument of the kind above that places gauge-fixed links near the identity needs plaquette bounds of order $o(n^{-2})$.

Up to the factor $L(b)^{1/2}$ the rate $b^{-1/2}$ of Proposition 75 is the right one, and the lower bound below holds for every n and every exterior.

Proposition 80. *Let $G = \text{SU}(2)$ with the chordal metric and $b = \beta_W > 0$. For every Gibbs specification of the gauge-fixed theory, every active link e and every configuration of the other links, the conditional law of Q_e satisfies*

$$\int \delta(1, Q_e) d\mu_e \geq \frac{1}{9 \max(6b, 1)^{1/2}}.$$

In particular $\varepsilon_V(\beta) \geq \frac{1}{9} \max(6b, 1)^{-1/2}$ for every n ; at every fixed n , $\varepsilon_V(\beta)$ lies between constant multiples of $b^{-1/2}$ and $(L(b)/b)^{1/2}$.

Proof. Write $\langle \cdot, \cdot \rangle$ for the Euclidean product of quaternions, so that $\text{Re Tr } q = 2\langle q, 1 \rangle$. The six plaquettes that contain e contribute $\exp(b\langle Q_e, W_p \rangle)$ to the conditional density, with $|W_p| = 1$; so μ_e has density $e^{\langle u, v \rangle} / Z(|v|)$ with respect to normalised Haar measure, for some v with $a = |v| \leq 6b$. Put $a' = \max(a, 1)$ and $\eta = a'^{-1/2} \leq 1$. On the geodesic ball of radius η about $v/|v|$ the exponent is at least $a(1 - \eta^2/2) \geq a - \frac{1}{2}$, and that ball has measure greater than $\eta^3/24$ (proof of Proposition 81); hence $Z(a) \geq e^{a-1/2} \eta^3/24$ and the density is at most $24e^{1/2} a^{3/2}$ (for $a < 1$ directly, since $Z \geq 1$ by Jensen's inequality). A chordal ball of radius r is a geodesic ball of radius θ with $2\sin(\theta/2) = r$, so $\theta \leq \pi r/2$, and its measure is $(\theta - \sin \theta \cos \theta)/\pi \leq 2\theta^3/(3\pi) \leq \pi^2 r^3/12$. With $r = \frac{2}{9} a'^{-1/2}$,

$$\mu_e(\delta(1, Q_e) < r) \leq 24e^{1/2} a^{3/2} \frac{\pi^2 r^3}{12} = 2e^{1/2} \pi^2 \left(\frac{2}{9}\right)^3 < \frac{1}{2},$$

so the mean of $\delta(1, Q_e)$ is at least $r/2 = \frac{1}{9} a'^{-1/2}$. Averaging over $e \in V$ and over the configuration gives the bound on ε_V . $\square$

Along the continuum tuning Theorem 71 says the following. If (D) holds, then in a cube of fixed physical size ℓ with the identity outside, the forest-gauge variables stay on average at distance at least $(1 - \alpha)D/3$ from the identity, and at least $1 - \alpha$ for $\text{SU}(2)$, at every spacing. This is not a contradiction: $n = \ell/a$ grows like $1/a$, while asymptotic freedom makes $\beta(a)$ grow only like $\log(1/a)$, and the corollary concerns fixed n . The size bound of Proposition 75 is likewise met along the tuning, since its right side grows only like $\log(1/a)$. It turns the hypothesis of Remark 61, that the gauge-fixed links decorrelate on the scale of the cube, into a necessary condition: (D) requires the forest-gauge variables to be disordered at a fixed physical scale. For gauge-invariant observables of a massive theory that disorder is clustering; for the gauge-fixed variables it is not proved.

The last statement concerns the following strategy: control the conditional law of one line of links, given everything else, uniformly in the configuration around it, and assemble the cube from lines. Fix a site x , let $s = \sigma(x)$ and $\mu \neq s$, and let $c_k = (x + ke_s, \mu)$, $k \in \mathbb{Z}$. All c_k are active, since σ is constant along the line through x .

Proposition 81. *Let $G = \text{SU}(2)$, identified with the unit quaternions, and $b = \beta_W \geq 1$. There is a configuration of the active links outside $\{c_k\}$, independent of b , for which the conditional law of $(Q_{c_k})_k$ has a ten-step transfer operator, similar to a*

positive self-adjoint operator K_b on $L^2(S^3)$ with simple top eigenvalue Λ_b , whose gap satisfies

$$0 < 1 - \frac{\Lambda_{b,1}}{\Lambda_b} \leq e^{20} 24^{10} b^{15} e^{-b/1280},$$

where $\Lambda_{b,1}$ is the next point of the spectrum.

Proof. The interactions. By Lemma 9(iii), c_k lies in two plaquettes of the plane $\{\mu, s\}$, which contain c_{k-1} or c_{k+1} , the link $(x + ke_s, s)$ or $(x + (k-1)e_s, s)$ of F , and an active twist link of direction s ; and in four plaquettes of the planes $\{\mu, t\}$ and $\{\mu, \rho\}$, each containing c_k , one link of F and two active links not on the line. The four plaquettes at c_k lie in the hyperplane of s -coordinate $x_s + k$, so these plaquettes at different k share no link; two plaquettes through c_k share no other link; and their links have directions other than s , so they are not twists. Set every twist to 1; then the plaquette at $x + ke_s$ in the plane $\{\mu, s\}$ contributes $b\langle q_k, q_{k+1} \rangle$ to the exponent, with $q_k = Q_{c_k}$. In each of the four other plaquettes at c_k set one of its two links to 1 and the other to X ; after a cyclic rotation and possibly an inversion the plaquette variable is $q_k X^{\pm 1}$ or $q_k^{-1} X^{\pm 1}$, and since $\text{Re}(q\bar{u}) = \text{Re}(\bar{q}u) = \langle q, u \rangle$ for unit quaternions, X can be chosen so that the plaquette contributes $b\langle q_k, u \rangle$ for any prescribed unit vector u . Every other active link is set to 1. Four unit vectors add up to any vector of length at most 4 (two copies each of $t \pm (1 - |t|^2)^{1/2} \mathbf{v}$ with $t = v/4$ and $\mathbf{v} \perp t$ a unit vector), so the conditional density of the line is proportional to $\exp(b \sum_k \langle q_k, q_{k+1} \rangle + \sum_k \langle h_k, q_k \rangle)$ for any fields h_k with $|h_k| \leq 4b$.

The fields. Let $a = \frac{1}{2}$, $c = \frac{\sqrt{3}}{2}$, let $e_1, \dots, e_4$ be the standard basis of $\mathbb{R}^4$, and let r_k have period ten with $(r_0, \dots, r_9) = a(e_1, -e_1, e_2, -e_2, e_3, -e_3, e_3, -e_2, e_2, -e_1)$. Put $h_k = b(2r_k - r_{k-1} - r_{k+1})$, so $|h_k| \leq 4ab = 2b$, and the configuration is independent of b because h_k/b is. With $v_k = q_k - r_k$ and $|q_k| = 1$, $|r_k| = a$, summation by parts gives, over $0 \leq k \leq 9$,

$$\begin{aligned} \frac{1}{2} \sum_k |v_{k+1} - v_k|^2 &= \sum_k (1 - \langle q_k, q_{k+1} \rangle) - \frac{1}{b} \sum_k \langle h_k, q_k \rangle \\ &\quad + \frac{1}{2} \sum_k |r_{k+1} - r_k|^2 + B(q_0) - B(q_{10}), \end{aligned}$$

with $B(q) = \langle q, r_0 - r_9 \rangle$. The boundary terms only conjugate the ten-step transfer kernel by a multiplication operator, so up to similarity it is

$$K_b(x, y) = \int_{(S^3)^9} \exp\left(-\frac{b}{2} \sum_{k=0}^9 |v_{k+1} - v_k|^2\right) dq_1 \cdots dq_9,$$

with $q_0 = x$, $q_{10} = y$ and normalised Haar measure. The sequence $r_0, \dots, r_{10}$ is a palindrome, so $K_b = B_b B_b^*$ for the five-step kernel B_b ; K_b is compact, positive and self-adjoint with a continuous positive kernel, its top eigenvalue Λ_b is simple, and its eigenfunction $\psi_b > 0$ is even under the reflection $x_4 \mapsto -x_4$, which commutes with K_b .

A barrier. Let $E = \frac{1}{2} \sum_{k=0}^9 |v_{k+1} - v_k|^2$ and suppose that $(q_0)_4 \geq 0 \geq (q_{10})_4$ and $E \leq \frac{1}{1280}$. Put $\Delta = (20E)^{1/2} \leq \frac{1}{8}$ and $w = v_0$; by Cauchy–Schwarz $|w - v_k| \leq \Delta$ for all k , and $0 \leq$

$w_4 \leq \Delta$ because the r_k have no fourth coordinate. From $|v_k + r_k| = 1$, $||w + r_k|^2 - 1| \leq \Delta(2 + \Delta) =: \eta$ for $0 \leq k \leq 5$. The pairs $r_k = \pm ae_l$ give $|w_l| \leq \eta/(2a) = \eta$ for $l = 1, 2, 3$ and $|w|^2 \geq \frac{3}{4} - \eta$, so $\frac{3}{4} \leq 3\eta^2 + \Delta^2 + \eta \leq 3(\frac{17}{64})^2 + \frac{1}{64} + \frac{17}{64} = \frac{2019}{4096}$, a contradiction. Hence $K_b(x, y) \leq e^{-b/1280}$ whenever $x_4 \geq 0 \geq y_4$.

The top eigenvalue. Let $\varepsilon = b^{-1/2}$. The geodesic ball of radius $\varepsilon \leq 1$ in S^3 has normalised measure $\frac{5}{\pi} \int_0^\varepsilon \sin^2 \theta d\theta \geq \frac{8\varepsilon^3}{3\pi^3} > \frac{\varepsilon^3}{24}$. If each q_k lies in the ball of radius ε about $r_k + ce_4$, then $|v_k - ce_4| \leq \varepsilon$ and $E \leq 20\varepsilon^2$, so the integrand is at least e^{-20} . Testing K_b on the normalised indicator of the ball about $r_0 + ce_4$ gives $\Lambda_b \geq e^{-20} (\varepsilon^3/24)^{10} = e^{-20} 24^{-10} b^{-15}$.

The gap. Let $\phi(x) = \text{sign}(x_4) \psi_b(x)$, a unit vector orthogonal to ψ_b . Then

$$\begin{aligned} \Lambda_b - \langle \phi, K_b \phi \rangle &= 4 \int_{x_4 > 0 > y_4} \psi_b(x) K_b(x, y) \psi_b(y) dx dy \\ &\leq e^{-b/1280}, \end{aligned}$$

because $\int_{x_4 > 0} \psi_b \leq \frac{1}{2}$ by Cauchy–Schwarz. By the variational principle $\Lambda_{b,1} \geq \langle \phi, K_b \phi \rangle$, and dividing by the lower bound on Λ_b proves the estimate; the gap is positive because Λ_b is simple. $\square$

The exterior of Proposition 81 is legal at every coupling, so a bound on the conditional law of single lines that is uniform in the exterior cannot be better than $e^{-b/1280}$ up to powers of b , and an argument of this kind cannot give (D) along the tuning. A proof of (D) has to use what such a bound discards: the joint resampling of the transverse links inside the cube, which the averaged coefficients $C_{V,j}$ retain, or the smallness of the probability of such exteriors under the Wilson measure.

XII. BLOCK AVERAGES

Proposition 63 locates the failure of (D) in the quadratic approximation: gauge transformations constant on the lines of F move the whole cube and change few links outside it. This section takes up the programme of the report [53] of the third author, which keeps the cube condition but changes the gauge and conditions the cube on block averages of the field, as in Bałaban’s renormalisation group [5, 6]. The natural block averages of the line field leave one class of lines invisible and fail at every size (Proposition 82). In a gauge whose fixed links are trees inside blocks of side two, the averages over whole blocks remove every zero mode, the constrained quadratic form has a gap that does not depend on the volume (Theorem 85), and the constrained cube condition holds at every coupling once the cube is large (Theorem 88). Integrating out one scale bounds the cube coefficients by the constrained ones plus those of the averages (Corollary 91), and the averages carry at least the coarse quadratic form at every scale (Proposition 89). For the group-valued links themselves, the straight paths between block roots satisfy a coercivity bound for all fields but do not compose, a continuous covariant map cannot treat the transverse positions alike, and a cut-off of Bałaban’s averaging

operation is coercive for all fields and agrees with the block averages to first order (Sec. XII B). It composes across all scales on fields of small curvature for every compact group (Sec. XII C), and under the SU(2) Wilson measure the fields that are not small form clusters of vanishing physical size (Sec. XII D).

A. Block-constrained cubes

In this subsection everything concerns the quadratic approximation: real fields φ on links, $(d\varphi)_p$ the signed sum around the plaquette p , and $(dh)_{(x,\mu)} = h(x) - h(x+e_\mu)$ for a function h on sites.

Proposition 82 ([53]). *In the setting of Proposition 63, condition the cube in addition on any family of the face averages*

$$V_{y,\mu}^{\text{face}}(\varphi) = \frac{1}{8} \sum_x (\varphi_{(x,\mu)} + \varphi_{(x+e_\mu,\mu)}),$$

where x runs over the eight sites of the block $2y + \{0,1\}^4$ with $x_\mu = 2y_\mu$. Then $\sum_{c \in C_I} K_c \geq 12N_n$ for the twelve classes

$$C_I = \{(0111, \mu), (1111, \mu), (0111 + e_\mu, \mu), (1111 + e_\mu, \mu) : \mu = 1, 2, 3\},$$

written as residue of the base and direction. Hence $K_c \geq N_n$ for some class, and the constrained condition fails for every $\alpha < 1$, at every even n and every b .

Proof. By Table III, $\sigma(0111) = \sigma(1111) = 0$, so the lines of F through the sites with x_1, x_2, x_3 odd have direction 0; call them invisible. Let ℓ be one of them and h_ℓ its indicator. The sum of dh_ℓ along a path of length two from x is $h_\ell(x) - h_\ell(x+2e_\mu)$. For $\mu \neq 0$ both end points have even μ -coordinate and are off ℓ ; for $\mu = 0$ they have the same transverse coordinates and are both on ℓ or both off it. So dh_ℓ changes no face average, and the translation argument of Proposition 63 applies to the constrained laws: $\sum_{i \in V} |(dh_\ell)_i| \leq \sum_{j \notin V} C_{V,j} |(dh_\ell)_j|$. An active link $e = (x, \mu)$ has $|(dh_\ell)_e| = 1$ for at most one invisible ℓ , since its two ends lie on different lines and both cannot be invisible (for $\mu \neq 0$ they differ in the coordinate μ , and for $\mu = 0$ the link would lie in F). It has one such ℓ exactly when its base or its end has residue 0111 or 1111, which by Table III are the classes of C_I : the six classes with base residue 0111 or 1111 and direction $\mu \neq 0$, and the six with base residue $r + e_\mu$, $r \in \{0111, 1111\}$, whose directions μ differ from $\sigma(r + e_\mu)$ (for instance $\sigma(0011) = 3 \neq 1$). Summing over the invisible lines gives $12N_n \leq \sum_{c \in C_I} K_c$. $\square$

Definition 83. The blocks are $B(y) = 2y + \{0,1\}^4$, $y \in \mathbb{Z}^4$, and $o = x \bmod 2$ is the offset of a site. A link (x, μ) is a *tree link* if $o_\mu = 0$ and $o_\nu = 0$ for every $\nu > \mu$, and *active* otherwise; fields vanish on tree links. For a coarse link (y, μ) the *block average* is

$$V_{y,\mu}(\varphi) = \frac{1}{16} \sum_{x \in B(y)} (\varphi_{(x,\mu)} + \varphi_{(x+e_\mu,\mu)}),$$

the mean of φ summed along the sixteen paths of length two in the direction μ that start in $B(y)$.

Lemma 84. (i) *The fifteen tree links of a block form a spanning tree of its sixteen sites. Each block carries 49 active links: the 17 active links inside it, with $o_\mu = 0$, and the 32 links with $o_\mu = 1$, which join $B(y)$ to $B(y+e_\mu)$.*

(ii) *If dh vanishes on the tree links, h is constant on every block, and $V_{y,\mu}(\varphi + dh) = V_{y,\mu}(\varphi) + h_{B(y)} - h_{B(y+e_\mu)}$.*

(iii) *With $S_y^\mu = \sum_{x \in B(y), o_\mu=0} \varphi_{(x,\mu)}$,*

$$V_{y,\mu}(\varphi) = \frac{1}{8} \sum_{x \in B(y), o_\mu=1} \varphi_{(x,\mu)} + \frac{1}{16} (S_y^\mu + S_{y+e_\mu}^\mu).$$

Proof. (i) A tree link of direction μ has $o_\mu = 0$ and stays in its block; there are 2^μ of them for each μ , fifteen in all. From any site, the tree links of direction 3 lead to $o_3 = 0$, then those of direction 2 to $o_2 = 0$, and so on to the corner, so the tree links connect the sixteen sites, and fifteen edges connecting sixteen vertices form a tree. Of the 64 links with base in the block, the 32 with $o_\mu = 0$ stay inside it. (ii) h is constant along the edges of a spanning tree. The path of length two from $x \in B(y)$ ends at $x+2e_\mu \in B(y+e_\mu)$, and the sum of dh along it is $h(x) - h(x+2e_\mu)$. (iii) The path from x consists of (x, μ) and $(x+e_\mu, \mu)$; a link with $o_\mu = 1$ is the second link of one path and the first of another, and the links of $S_{y+e_\mu}^\mu$ are the second links of the paths that start with $o_\mu = 1$. $\square$

By (ii) the residual gauge freedom is the coarse gauge group, which acts on the block averages as on the links of the coarse lattice. Let W be a set of blocks and let Φ_W be the space of finitely supported real φ on the active links with base in $\bigcup W$, with $V_{y,\mu}(\varphi) = 0$ for every $y \in W$ and every μ . Write $Q(\varphi) = \sum_p (d\varphi)_p^2$ over all plaquettes, φ being zero off its support.

Theorem 85. *For every set of blocks W and every $\varphi \in \Phi_W$,*

$$\sum_e \varphi_e^2 \leq 52 Q(\varphi).$$

Proof. Inside a block. Let $e = (x, \mu)$ be an active link with $o_\mu = 0$, and let $v_1 > \dots > v_k$ be the directions above μ in which o equals one; $1 \leq k \leq 3$. Put $x_0 = x$ and $x_i = x_{i-1} - e_{v_i}$. The plaquette p_i at x_i in the plane $\{\mu, v_i\}$ has the links (x_i, μ) , (x_{i-1}, μ) and the tree links (x_i, v_i) , $(x_i + e_\mu, v_i)$, whose offsets vanish in v_i and above it; and (x_k, μ) is a tree link. Hence $\varphi_e = \pm(d\varphi)_{p_1} \pm \dots \pm (d\varphi)_{p_k}$ and $\varphi_e^2 \leq k \sum_i (d\varphi)_{p_i}^2$. The plaquette at x' in the plane $\{\mu, v\}$ is used only by the links $(x' + e_\nu + \sum_{\rho \in R} e_\rho, \mu)$ with $R \subset \{v+1, \dots, 3\}$, each with $k = c+1 + |R|$, where c is the number of ones of the offset of x' strictly between μ and v . Its total weight $2^{3-v}(c+1) + (3-v)2^{2-v}$ is at most 8, the value at $\mu = 0$, $v = 1$. So the sum I of φ_e^2 over the active links inside the blocks is at most $8Q_{\text{in}}$, where Q_{in} is the sum over the plaquettes inside a block.

Between blocks. Fix $y \in W$ and μ , and label the eight links (x, μ) , $x \in B(y)$, $o_\mu = 1$, by the transverse offset $a \in \{0,1\}^3$. If a and a' differ in the coordinate v , with $o_v = 0$ at a , the plaquette at x in the plane $\{\mu, v\}$ contains these two links,

(x, v) in $B(y)$ and $(x + e_\mu, v)$ in $B(y + e_\mu)$, the last two inside their blocks. So $\varphi_a - \varphi_{a'} = \pm(d\varphi)_p \pm \psi \pm \psi'$ with ψ, ψ' inside blocks, and these twelve face plaquettes correspond to the twelve edges of the cube $\{0, 1\}^3$. By Lemma 84(iii) and $V_{y,\mu}(\varphi) = 0$, the mean of the φ_a is $m = -(S_y^\mu + S_{y+e_\mu}^\mu)/16$, so $8m^2 \leq \frac{1}{2}$ times the sum of the squares of the sixteen links of S_y^μ and $S_{y+e_\mu}^\mu$. The Laplacian of the graph $\{0, 1\}^3$ has smallest nonzero eigenvalue 2, so

$$\begin{aligned} \sum_a \varphi_a^2 &= 8m^2 + \sum_a (\varphi_a - m)^2 \\ &\leq 8m^2 + \frac{3}{2} \sum_{\text{edges}} ((d\varphi)_p^2 + \psi^2 + \psi'^2). \end{aligned}$$

Summed over $y \in W$ and μ , a link (z, v) inside a block occurs as ψ or ψ' once for each $\mu \neq v$ (as ψ if $o_\mu = 1$, as ψ' of the face of the block below if $o_\mu = 0$) and in two of the sums S^v , so it receives weight at most $3 \cdot \frac{3}{2} + 2 \cdot \frac{1}{2} = \frac{11}{2}$. The face plaquettes are distinct and lie in no block, so the links between blocks contribute at most $\frac{3}{2}Q_{\text{face}} + \frac{1}{2}I$, and

$$\sum_e \varphi_e^2 \leq \frac{13}{2}I + \frac{3}{2}Q_{\text{face}} \leq 52Q_{\text{in}} + \frac{3}{2}Q_{\text{face}} \leq 52Q(\varphi). \quad \square$$

With W all blocks the bound extends to ℓ^2 , so the block-constrained quadratic form on $\mathbb{Z}^4$ has a spectral gap at least $\frac{1}{52}$; it removes in particular the four flat connections that are zero modes of the unconstrained form. In the Bloch decomposition the exact gap is $0.10748776948684\dots$, attained at the momentum $(0, \pi, \pi, \pi)$, a floating-point value computed in [53] and reproduced in the archive; the theorem does not use it. The cube conditions now take the following form. Let n be even, V the active links with base in $[0, n]^4$, W its blocks, and for an exterior η and values $v = (v_{y,\mu})_{y \in W}$ let $\gamma_V^\epsilon(\cdot | \eta, v)$ be the law on $\mathbb{R}^V$ with density proportional to $\exp(-\frac{b}{2} \sum_{p \cap V \neq \emptyset} (d\varphi)_p^2)$ on the affine set $\{V_{y,\mu}(\varphi) = v_{y,\mu}\}$, with $\varphi = \eta$ off V . Define $C_{V,j}^\epsilon$, K_c^ϵ and $D_c(n, \alpha)$ as $C_{V,j}$, K_c and $D(n, \alpha)$, with v held fixed and the 49 classes of active links modulo two.

Lemma 86. *Let $H \geq 0$ be symmetric on $\mathbb{R}^m$ with $\|H\| \leq h$, let $B : \mathbb{R}^m \rightarrow \mathbb{R}^k$ with $BB^\top \geq \beta > 0$ and $\langle x, Hx \rangle \geq a|x|^2$ on $\ker B$, and $s > 0$. Then $K = \begin{pmatrix} H & sB^\top \\ sB & 0 \end{pmatrix}$ is invertible, and the solution of $K(x, y) = (f, g)$ has $|x| \leq (|f| + h|x_1|)/a + |x_1|$ with $|x_1| \leq |g|/(s\sqrt{\beta})$, and $|y| \leq (|f| + h|x|)/(s\sqrt{\beta})$.*

Proof. Put $x_1 = B^\top(BB^\top)^{-1}g/s$, so $sBx_1 = g$ and $|x_1| \leq |g|/(s\sqrt{\beta})$. On $\ker B$ the compression of H is at least a , so there is a unique $x_0 \in \ker B$ with $Hx_0 - f + Hx_1 \perp \ker B$, and $a|x_0| \leq |f| + h|x_1|$. Then $f - H(x_0 + x_1)$ lies in $(\ker B)^\perp = \text{ran } B^\top$, and $y = (BB^\top)^{-1}B(f - Hx)/s$ solves $sB^\top y = f - Hx$. For $f = g = 0$ the construction gives zero, so K is injective. $\square$

Lemma 87 (Combes–Thomas [27]). *Let K be an invertible symmetric matrix indexed by a finite set X , with $\|K^{-1}\| \leq M_0$ and $\sum_b |K_{ab}| \leq S$ for every a , and let $d : X \rightarrow \mathbb{R}$ satisfy $|d(a) - d(b)| \leq 1$ whenever $K_{ab} \neq 0$. If $(e^\lambda - 1)SM_0 \leq \frac{1}{2}$, then $\|e^{\lambda d} K^{-1} e^{-\lambda d}\| \leq 2M_0$.*

Proof. $e^{\lambda d} K e^{-\lambda d} = K + E$ with $|E_{ab}| \leq (e^\lambda - 1)|K_{ab}|$, so $\|E\| \leq (e^\lambda - 1)S$ by the Schur test, $\|K^{-1}E\| \leq \frac{1}{2}$, and $(K + E)^{-1} = (1 + K^{-1}E)^{-1}K^{-1}$ has norm at most $2M_0$. $\square$

Theorem 88. *For every $b > 0$, every even n and every exterior link j :*

- (i) $\gamma_V^\epsilon(\cdot | \eta, v)$ is a Gaussian law whose covariance depends on neither η nor v and whose mean is affine in (η, v) , and $C_{V,j}^\epsilon = \|u_j\|_1$, where $u_j \in \mathbb{R}^V$ is the change of the mean when η_j increases by one; u_j does not depend on b ;
- (ii) $\sum_{e \in V} e^{2\lambda_0 d(e)} u_j(e)^2 \leq 3500^2$ and $\|u_j\|_1 \leq 10^{14}$, where $\lambda_0 = 2.4 \cdot 10^{-5}$ and $d(e)$ is the ℓ^∞ distance between the midpoints of e and j ;
- (iii) $D_c(n, \alpha)$ holds for every $\alpha \in (0, 1)$ and every even $n \geq 2 \cdot 10^{15}/\alpha$, at every $b > 0$.

Proof. (i) Let H be the matrix of $\sum_{p \cap V \neq \emptyset} (d\varphi)_p^2$ on $\mathbb{R}^V$ and B that of the averages $V_{y,\mu}$, $y \in W$, restricted to V . By Theorem 85 with this W , $H \geq \frac{1}{52}$ on $\ker B$, so the density is that of a non-degenerate Gaussian on the affine set, its covariance is b^{-1} times the inverse of H on $\ker B$, and its mean m minimises the quadratic form on the affine set. The mean is affine in (η, v) , and two exteriors that differ at j give translates by u_j , which minimises $Q(u + \delta_j)$ subject to $Bu = g$, g being minus the coefficients of j in the averages. The translation argument of Proposition 63 gives $C_{V,j}^\epsilon = \|u_j\|_1$.

(ii) The pair (u_j, y) , with y the Lagrange multipliers, solves $K(u_j, y) = (f, sg)$ with K as in Lemma 86, $s = 96$ and $f = -\sum_{p \ni j} (d\delta_j)_p d_p|_V$, where $d_p|_V$ is the row of the plaquette p restricted to V . Each link lies in six plaquettes of four links, so the row sums of $|H|$ are at most 24 and $h \leq 24$; $|f| \leq 6\sqrt{3}$; and $|g| \leq \frac{1}{16}$, since an exterior link enters at most one average of W , with weight $\frac{1}{16}$. By Lemma 84(iii) each average has squared norm at least $8 \cdot \frac{1}{64} = \frac{1}{8}$, two averages share links only if they are (y, μ) and $(y \pm e_\mu, \mu)$, and then their inner product is at most $8 \cdot \frac{1}{256} = \frac{1}{32}$; so $BB^\top \geq \frac{1}{16}$ by Gershgorin's theorem. With $a = \frac{1}{52}$, Lemma 86 gives $\|K^{-1}\| \leq M_0 = 105$. The rows of K have absolute sums at most $S = \max(24 + \frac{96}{8}, 96 \cdot 2) = 192$, because an average has eight coefficients $\frac{1}{8}$ and sixteen of $\frac{1}{16}$ and a link enters averages with total weight at most $\frac{1}{8}$. Index the rows of B by the centre $2y + \frac{1}{2}(1, 1, 1, 1) + e_\mu$ of the two blocks of the average and put d equal to the ℓ^∞ distance to the midpoint of j . Links of one plaquette, and an average and its links, are within distance one, so Lemma 87 applies with $\lambda = \log(1 + 1/40320) > \lambda_0$, and

$$\sum_e e^{2\lambda d(e)} u_j(e)^2 \leq (2M_0 e^\lambda (|f| + s|g|))^2 \leq 3500^2,$$

since f and g live at distance at most one from j . At most $4(64k^3 + 16k)$ links have their base at ℓ^∞ distance $k \geq 1$ from the base of j , and their midpoints are at distance at least $k - 1$, so Cauchy–Schwarz gives

$$\|u_j\|_1^2 \leq 3500^2 \cdot 4e^{2\lambda} \left(1 + \sum_{k \geq 1} (64k^3 + 16k)e^{-2\lambda k}\right) \leq (10^{14})^2.$$

(iii) $C_{V_x, j}^c = 0$ unless j shares a plaquette with V_x or enters an average of V_x , which forces the base z of j into $x + [-1, n+1]^4$, while $z \notin x + [0, n]^4$ when $j \notin V_x$. For each coordinate at most $\frac{n}{2} + 2$ even x_i satisfy the first condition and exactly $\frac{n}{2}$ the second, so

$$K_c^c \leq \left(\left(\frac{n}{2} + 2 \right)^4 - \left(\frac{n}{2} \right)^4 \right) 10^{14} \leq \frac{17}{n} 10^{14} N_n$$

for $n \geq 10^4$, which is at most αN_n when $n \geq 2 \cdot 10^{15} / \alpha$. $\square$

The constants are crude. The computations of [53] give $C_{V, j}^c \leq 12.48$ at $n = 4$, and an exact face expansion $K_c^c / N_n = a_c / n + O(n^{-2})$ with $\max_c a_c = 61.18$, so that $D_c(n, 0.95)$ first holds at $n = 64$; they are floating-point values, reproduced in the archive for $n \leq 4$, and the theorem does not use them. The coefficients are large because the constrained responses obey no maximum principle: an exterior link changes the curvature of up to six plaquettes, and keeping the curvature small moves a dozen links inside the cube with alternating signs. Re-weighting the classes does not help much: if $\sum_{j \in c \setminus V} \sum_{c'} w_{c'} C_{V, j}^{c'} \leq \alpha N_n w_c$ for positive weights w , where $C_{V, j}^{c'}$ is the part of $\|u_j\|_1$ on the links of the class c' , then αN_n is at least the spectral radius of the non-negative matrix of these sums, hence at least $\min_c K_c^c$, by the Collatz–Wielandt inequality [53].

The block averages can be integrated out one scale at a time. For a field v on the coarse links put $Q_1(v) = \sum_P (dv)_P^2$ over the plaquettes P of the coarse lattice, and write $V(\varphi) = (V_{y, \mu}(\varphi))_{y, \mu}$. The averages of averages are formed in the same way on the coarse lattice, and $V^{(k)}$, Q_k denote the k -fold averages and the plaquette form of the lattice of spacing 2^k .

Proposition 89. *Let φ be a finitely supported field on $\mathbb{Z}^4$ that vanishes on the tree links.*

- (i) *For the coarse plaquette P at y in the plane $\{\mu, \nu\}$, $16(dV(\varphi))_P$ is the sum over $x \in B(y)$ of the curvatures of the squares of side two spanned by $x, x + 2e_\mu, x + 2e_\nu$, and every plaquette of $\mathbb{Z}^4$ lies in exactly four of these squares. Hence $Q_1(V(\varphi)) \leq Q(\varphi)$, and $Q_k(V^{(k)}(\varphi)) \leq Q(\varphi)$ for every $k \geq 1$.*
- (ii) *For every finitely supported v on the coarse links, the field $\varphi^*(v)$ equal to $v_{y, \mu}$ on the eight links (x, μ) , $x \in B(y)$, $o_\mu = 1$, and to zero on the other links, vanishes on the tree links, has $V(\varphi^*(v)) = v$ and $Q(\varphi^*(v)) = 4Q_1(v)$.*
- (iii) *The effective form $S_1(v) = \inf\{Q(\varphi) : V(\varphi) = v\}$ satisfies $Q_1 \leq S_1 \leq 4Q_1$, and $S_k(w) = \inf\{Q(\varphi) : V^{(k)}(\varphi) = w\} \geq Q_k(w)$ for every k , the infima over finitely supported fields that vanish on the tree links.*
- (iv) *For every k , every set W of blocks of the lattice of spacing 2^k and every w in the space Φ_W of that lattice, $\sum_e w_e^2 \leq 52S_k(w)$.*

Proof. (i) The path of length two from $x \in B(y)$ in the direction μ ends at $x + 2e_\mu \in B(y + e_\mu)$, at the same offset, and so on around the square; so the four averages of $(dV)_P$

pair the sixteen paths into sixteen closed squares, and the sum of φ around a square is the sum of the curvatures of its four plaquettes. A plaquette at z in the plane $\{\mu, \nu\}$ lies in the square from x exactly when $z - x \in \{0, 1\}e_\mu + \{0, 1\}e_\nu$, which gives four points x , each in one block. By Cauchy–Schwarz, $(dV)_P^2 \leq \frac{1}{16} \sum_x \ell_x^2 \leq \frac{1}{4} \sum_x \sum_{P \subset \square_x} (d\varphi)_P^2$, where ℓ_x is the curvature of the square from x ; summing over P gives $Q_1(V(\varphi)) \leq Q(\varphi)$. The argument uses neither the tree links nor the spacing, so it applies on every lattice, and induction gives the bound for $V^{(k)}$.

(ii) The tree links have $o_\mu = 0$. By Lemma 84(iii) the average of φ^* is $\frac{1}{8} \cdot 8v_{y, \mu}$. A plaquette inside a block has only links with $o_\mu = 0$ in its direction μ , and so vanishes. A plaquette at x in the plane $\{\mu, \nu\}$ with $o_\mu = 1, o_\nu = 0$ has the two links (x, μ) , $(x + e_\nu, \mu)$ of the same block, both equal to $v_{y, \mu}$, and two links with offset zero in the direction ν ; its curvature vanishes, and so does that of the plaquettes with $o_\mu = 0, o_\nu = 1$. If $o_\mu = o_\nu = 1$, its four links carry $v_{y, \mu}, v_{y+e_\mu, \nu}, v_{y+e_\nu, \mu}, v_{y, \nu}$, so its curvature is $(dv)_P$, and there are four such plaquettes for each P .

(iii) The lower bound is (i) and the upper bound is (ii); for S_k apply (i) to the k -fold averages.

(iv) Theorem 85 holds on every lattice, so $\sum_e w_e^2 \leq 52Q_k(w) \leq 52S_k(w)$ by (iii). $\square$

In the Bloch decomposition the eigenvalues of S_1 relative to Q_1 fill $[1, 4]$: they tend to one as the momentum tends to zero and equal four at $(\pi, \pi, 0, 0)$, a floating-point computation reproduced in the archive. So at this order the averages carry the coarse quadratic form with no loss at long wavelengths; the coupling does not run here, and its running is the order- g^2 effect of Remark 92(c). Part (iv) is a gap for the averages at every scale, with the constant of Theorem 85; it does not give the localisation of Theorem 88(ii) for S_k , since S_k is not a nearest-neighbour form.

Lemma 90. *Let (X, d_X) and (Y, d_Y) be Polish spaces, ρ, ρ' Borel probability measures on Y , and κ, κ' probability kernels from Y to X , all with finite first moments. If $W_1(\kappa(\cdot | v), \kappa'(\cdot | v)) \leq \varepsilon$ and $W_1(\kappa'(\cdot | v), \kappa'(\cdot | v')) \leq d_Y(v, v')$ for all v, v' , then the mixtures $\pi = \int \kappa(\cdot | v) \rho(dv)$ and $\pi' = \int \kappa'(\cdot | v) \rho'(dv)$ satisfy*

$$W_1(\pi, \pi') \leq \varepsilon + W_1^{d_Y}(\rho, \rho').$$

Proof. Let F be 1-Lipschitz on X and $\kappa F(v) = \int F d\kappa(\cdot | v)$. Then $\int F d\pi - \int F d\pi' = \int (\kappa F - \kappa' F) d\rho + (\int \kappa' F d\rho - \int \kappa' F d\rho')$. The first term is at most ε . The function $\kappa' F$ is 1-Lipschitz for d_Y , so the second is at most $W_1^{d_Y}(\rho, \rho')$. The Kantorovich–Rubinstein duality gives the claim. $\square$

A law γ on $\mathbb{R}^V$ decomposes along the averages if $\gamma(\cdot | \eta) = \int \gamma_V(\cdot | \eta, v) \rho_V(dv | \eta)$, with $\rho_V(\cdot | \eta)$ the law of $(V_{y, \mu})_{y \in W}$. Every law with density proportional to $\exp(-\frac{b}{2} \sum_{P \cap V \neq \emptyset} (d\varphi)_P^2) g(V(\varphi))$, on a set cut out by conditions on the averages, decomposes in this way, because Lebesgue measure disintegrates along the fibres of a surjective linear map into multiples of Lebesgue measure on the fibres. Here g can carry any action of the averages, in particular a coarse gauge fixing and a coarse effective action, which is the form a law takes after the fine scale has been integrated out.

Corollary 91. Let $r_w \in \mathbb{R}^V$ minimise $\sum_{p \cap V \neq \emptyset} (d\phi)_p^2$, with zero exterior, subject to $V_{y', \mu'}(r_w) = \delta_{w, (y', \mu')}$ for all (y', μ') , $y' \in W$; put $L_w = \|r_w\|_1$ and $d_L(v, v') = \sum_w L_w |v_w - v'_w|$. For every law that decomposes along the averages,

$$C_{V,j} \leq C_{V,j}^c + C_{V,j}^p, \quad C_{V,j}^p = \sup \frac{W_1^{d_L}(\rho_V(\cdot | \eta), \rho_V(\cdot | \eta'))}{|\eta_j - \eta'_j|},$$

the supremum over exteriors that differ only at j , and $L_w \leq 4 \cdot 10^{14}$. Hence $K_c \leq K_c^c + K_c^p$ with $K_c^p = \sum_{j \in c \setminus V} C_{V,j}^p$, and if $K_c^p \leq \frac{\alpha}{2} N_n$ for every class and $n \geq 4 \cdot 10^{15}/\alpha$, then $D(n, \alpha)$ holds for the law, with the 49 classes of active links.

Proof. Apply Lemma 90 with $\kappa = \gamma_V^c(\cdot | \eta, \cdot)$, $\kappa' = \gamma_V^c(\cdot | \eta', \cdot)$, $\rho = \rho_V(\cdot | \eta)$, $\rho' = \rho_V(\cdot | \eta')$. The first condition holds with $\varepsilon = C_{V,j}^c |\eta_j - \eta'_j|$. By Theorem 88(i) the laws $\gamma_V^c(\cdot | \eta', v)$ and $\gamma_V^c(\cdot | \eta', v')$ are translates by $\sum_w (v'_w - v_w) r_w$, so their distance is at most $d_L(v, v')$. For the bound on L_w , the pair (r_w, y) solves $K(r_w, y) = (0, s e_w)$ with K as in the proof of Theorem 88. Measure d from the centre of the row w ; Lemma 87 gives $\sum_e e^{2\lambda d(e)} r_w(e)^2 \leq (2M_0 s)^2 = 20160^2$. With $z_0 = 2y + e_\mu$ for $w = (y, \mu)$, a link whose base is at distance k from z_0 has its midpoint at distance at least $k - 1$ from the centre, and the count in that proof gives $L_w^2 \leq 20160^2 \cdot 4e^{2\lambda} (1 + \sum_{k \geq 1} (64k^3 + 16k)e^{-2\lambda k}) \leq (4 \cdot 10^{14})^2$. Summing over the translates as in Definition 42 gives $K_c \leq K_c^c + K_c^p$, and Theorem 88(iii) with $\alpha/2$ in place of α gives $K_c^c \leq \frac{\alpha}{2} N_n$. $\square$

The responses are much smaller than the bound: in floating point, $\max_w L_w$ is 23.8, 44.1 and 50.9 for the block (m, m, m, m) , $m = \lfloor n/4 \rfloor$, of the cubes $n = 4, 6, 8$. Since Theorem 88 controls the first term at every coupling, the corollary reduces $D(n, \alpha)$ for a law that decomposes along the averages to the cube condition K_c^p for its averages, a problem of the same form one scale up, in the metric d_L . Iterating it needs the localisation of Theorem 88(ii) for S_k and ends in a cube condition for the coarsest averages, which is where a gap has to come from: the quadratic approximation describes a massless field and has none. The corollary is one step of the composition of Remark 92(a), not its conclusion.

Remark 92. Theorem 88 is the weak-coupling half of the constrained route at leading order, and Theorem 45 cannot use it as it stands. Four steps remain, each an open problem.

- (a) *Composition.* D_c holds the block averages fixed. A gap in physical units needs a theorem that combines the constrained conditions at all scales with a cube condition for the coarsest averages, in the place of Theorem 45. Corollary 91 is one step of it, and Proposition 89(iv) gives the gap at every scale; the localisation of S_k uniformly in k and a cube condition for the coarsest averages are open.
- (b) *Beyond the quadratic approximation.* For the Wilson measure the block averages are products of group elements along paths, their level sets are curved, and the Bakry-Émery tensor of Sec. VI is negative at large fields; Theorem 85 controls only the small fields, and the large fields have to be split off as in Balaban's analysis [7]. For the

straight paths between block roots the group-valued estimate holds for all fields (Theorem 94), but that map does not compose across scales (Proposition 97), and for every compact connected group no continuous covariant map treats the transverse positions of a block alike (Proposition 98). The cut-off average of Theorem 100 is coercive for all fields and agrees with the averages to first order, with constants for one step; on torus-valued fields of small curvature it composes (Theorem 102), and for every compact group it composes on fields whose curvature is small at every scale with a small sum (Theorem 106); under the Wilson measure the remaining large fields form microscopic clusters (Corollary 109), but the conditional estimate needed near them is open.

- (c) *Second order.* Let $b_0 = 11N/(48\pi^2)$ for $SU(N)$, $t = 2b_0 \log 2$, and let $g_m^{-2} = g_*^{-2} + tm$ be the one-loop coupling m steps above a matching scale. Comparison with integrals gives $\sum_{m \leq M} g_m^2 = t^{-1} \log M + O(1)$ and $\sum_m g_m^4 \leq \pi^2/(6t^2)$ [53]. A loss $1 + c g_m^2$ per step therefore compounds to a power of M , and a step from one scale to the next has to be exact at order g^2 , with that term absorbed into the running coupling and the normalisation of the field, and remainders of order g^4 .
- (d) *Crossover and existence.* The passage from weak coupling to the strong-coupling range of Sec. IV–VIII and the existence of the limit, (E), are untouched.

The renormalisation group needed for (a) and (b) is the ultraviolet half of (E), so the constrained route places both residues on one construction.

B. Path blocks for the group-valued links

Remark 92(b) asks for estimates on the actual group-valued links. They can be proved for a coarser block map, the straight path between block roots, and this map exhibits the second half of the difficulty: it does not compose across scales. A topological argument shows why no continuous map that treats the transverse positions alike can take its place.

Let $G \subset U(r)$ be a compact matrix group, $\delta(A, B) = \|A - B\|_{HS}$, which is invariant under left and right multiplication, and work on the periodic lattice of side $2L$ with the block tree of Definition 83, all fifteen tree links of each block equal to I . Write $\mathcal{B}$ for the number of blocks. The *path map* is

$$V_{y, \mu}^p(U) = U_{2y + e_\mu, \mu},$$

the product along the straight path from $2y$ to $2y + 2e_\mu$, whose first link is a tree link; it transforms as a link under the gauge transformations at the block roots. Fixing $V^p = v$ fixes four of the 49 active links of each block, so the fibre is $G^{45\mathcal{B}}$, and the Wilson law in the tree gauge restricts to it with product Haar measure as reference, since the tree gauge has unit Jacobian. The *reconstruction* $R(v)$ puts $R_{2y+o, \mu} = v_{y, \mu}$ if $o_\mu = 1$ and $R_{2y+o, \mu} = I$ otherwise. For U in the fibre of v put

$$D_v(U)^2 = \sum_e \delta(U_e, R_e(v))^2, \quad Q(U) = \sum_p \delta(U_p, I)^2,$$

the first sum having zero terms on the tree and path links, and let $Q_1(v) = \sum_P \delta(v_P, I)^2$ on the coarse lattice.

Lemma 93. *For vectors z_a of a real Hilbert space indexed by $a \in \{0, 1\}^3$, with $z_0 = 0$,*

$$\sum_a \|z_a\|^2 \leq \frac{3 + \sqrt{7}}{2} \sum_{\{a, a'\}} \|z_a - z_{a'}\|^2 \leq 3 \sum_{\{a, a'\}} \|z_a - z_{a'}\|^2,$$

the sums on the right over the twelve edges of the cube.

Proof. It suffices to treat real numbers. The right side is $\langle z, L_0 z \rangle$, where L_0 is the Laplacian of the cube with the vertex 0 deleted, and the claim is that its least eigenvalue is $3 - \sqrt{7}$. The permutations of the coordinates commute with L_0 . On the functions f_k of the level $k = |a|$, L_0 acts by $f_1 \mapsto 3f_1 - 2f_2$, $f_2 \mapsto 3f_2 - 2f_1 - f_3$, $f_3 \mapsto 3f_3 - 3f_2$, whose characteristic polynomial in $t = 3 - \lambda$ is $t^3 - 7t$; the eigenvalues are 3 and $3 \pm \sqrt{7}$. On the functions of levels one and two whose sum over each level vanishes, indexed by the coordinate that is one at level one and zero at level two, L_0 acts by the matrix $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, with eigenvalues 2 and 4 ; at level three these functions vanish. The two spaces have dimensions 3 and 4 and span $\mathbb{R}^7$. Finally $1/(3 - \sqrt{7}) = (3 + \sqrt{7})/2 < 3$.¹⁸ $\square$

Theorem 94. *For every $G \subset U(r)$, every $L \geq 1$, every v and every U in the fibre of v ,*

$$D_v(U)^2 \leq 224Q(U).$$

Proof. Let I_0 be the sum of $\delta(U_e, I)^2$ over the active links inside the blocks and J_0 the sum of $\delta(U_e, v_{y,\mu})^2$ over the links that leave $B(y)$ in the direction μ , so $D_v^2 = I_0 + J_0$. For an active link inside a block, the chain of plaquettes in the proof of Theorem 85 applies to group elements: each plaquette of the chain has its two transverse links on the tree, so the holonomy is the product of two consecutive links of the chain, and invariance of δ with the triangle inequality gives $\delta(U_e, I)^2 \leq k \sum_i \delta(U_{p_i}, I)^2$ with $k \leq 3$. The loads are those of that proof, so $I_0 \leq 8Q_{\text{in}}$. Fix y, μ and label the eight links leaving $B(y)$ by the transverse offset $a \in \{0, 1\}^3$; the link $a = 0$ is the path link and equals $v_{y,\mu}$. For a transverse edge $\{a, a'\}$ the face plaquette has holonomy $U_a \psi' U_{a'}^{-1} \psi^{-1}$ with ψ, ψ' links inside the two blocks, so $\delta(U_a, U_{a'}) \leq \delta(U_p, I) + \delta(\psi, I) + \delta(\psi', I)$. Lemma 93 with $z_a = U_a - v_{y,\mu}$ in the real Hilbert–Schmidt space and Cauchy–Schwarz bound the eight terms by 9 times the sum of the three squares over the twelve edges. Each face plaquette occurs once, and each link inside a block occurs once for each of the three directions other than its own, as in the proof of Theorem 85. Hence $J_0 \leq 9Q_{\text{face}} + 27I_0$ and $D_v^2 \leq 28I_0 + 9Q_{\text{face}} \leq 224Q$. $\square$

The bound holds for all fields, with no small-field condition and no logarithmic chart; Theorem 85 is its linearisation for the averages, with the better constant 52.

Proposition 95. $Q_1(V^p(U)) \leq 4Q(U)$, $Q(R(v)) = 4Q_1(v)$ and $Q(U) \leq 8Q_1(v) + 48D_v(U)^2$ for U in the fibre of v ; if $Q_1(v) = 0$, then $Q(U) \leq 24D_v(U)^2$.

Proof. The holonomy of a coarse plaquette is that of the square of side two at $2y$, a product of four conjugated fine plaquette holonomies, so $\delta(v_P, I) \leq \sum_{p \subset P} \delta(U_p, I)$; these squares tile each coordinate plane, which gives the first bound. In $R(v)$ only the plaquettes with offset one in both directions carry holonomy, equal to that of the coarse plaquette, four for each P . For any plaquette, telescoping gives $\delta(U_p, R_p) \leq \sum_{e \in p} \delta(U_e, R_e)$, so $\delta(U_p, I)^2 \leq 2\delta(R_p, I)^2 + 8 \sum_{e \in p} \delta(U_e, R_e)^2$, and each link lies in six plaquettes. If $Q_1(v) = 0$, every R_p is I and the factor 2 is not needed. $\square$

Let $m = 45\mathcal{B}$, let the action be λQ with $\lambda \geq 0$ (for $SU(N)$, $\lambda = \beta/(2N)$ up to a constant), and put

$$z_G(t) = \int_G e^{-t\delta(g, I)^2} dg, \quad Z_\lambda(v) = \int_{G^m} e^{-\lambda Q(U)} dU,$$

with normalised Haar measures, the integral over the fibre of v .

Theorem 96. *For every v ,*

$$e^{-8\lambda Q_1(v)} z_G(48\lambda)^m \leq Z_\lambda(v) \leq e^{-\lambda Q_1(v)/8} z_G(\lambda/448)^m,$$

and if $Q_1(v) = 0$, then $z_G(24\lambda)^m \leq Z_\lambda(v) \leq z_G(\lambda/224)^m$. Let $A_\lambda = z_G(\lambda/448)/z_G(24\lambda)$. Under the conditional law at such a v ,

$$\mathbb{E} e^{\lambda D_v^2/448} \leq A_\lambda^m, \quad \mathbb{E} \delta(U_e, I)^2 \leq \frac{20160}{\lambda} \log A_\lambda$$

for each link e when $v \equiv I$, and $A_\lambda \rightarrow 10752^{q/2}$ as $\lambda \rightarrow \infty$ if G has dimension $q \geq 1$.

Proof. By Theorem 94 and Proposition 95, $Q \geq \frac{1}{2}(D_v^2/224 + Q_1/4)$ and $Q \leq 8Q_1 + 48D_v^2$. Each term of D_v^2 involves one free link, and right multiplication by $R_e(v)^{-1}$ preserves Haar measure, so $\int e^{-tD_v^2} dU = z_G(t)^m$; this gives the two bounds, and the flat case is the same with $Q \leq 24D_v^2$. Next, $-\lambda Q + \lambda D_v^2/448 \leq -\lambda D_v^2/448$, and integration with the lower bound gives the exponential moment; Jensen's inequality gives $\mathbb{E} D_v^2 \leq (448m/\lambda) \log A_\lambda$. When $v \equiv I$ the law is invariant under translation by blocks, so the $\mathcal{B}$ links of each of the 45 classes have equal expectations, and $448 \cdot 45 = 20160$. Finally, in exponential coordinates $\delta(e^X, I)^2 = \|X\|^2 + O(\|X\|^4)$ and the Haar density is smooth and positive at 0, while $\delta(\cdot, I)$ is bounded below away from I ; so $z_G(t) \sim c_G t^{-q/2}$, and the ratio at $\lambda/448$ and 24λ tends to $(24 \cdot 448)^{q/2}$. $\square$

So the fluctuations of the free links around a flat path field are of order $\lambda^{-1/2}$ uniformly in the volume, for the group variables themselves. The path map cannot be iterated, however, and this is where it differs from the averages. Write $V^{p,k}$ for the k -fold path map, the straight path of length $s = 2^k$, and Q_k for the plaquette energy of the lattice of spacing s .

¹⁸ The constant $(3 + \sqrt{7})/2$ is attained by a function of the level $|a|$ alone, so it is the best constant in the lemma.

Proposition 97. *On the periodic lattice of side $3s$, $s = 2^k$, there are $SU(2)$ fields in the tree gauge with*

$$\frac{Q(U)}{Q_k(V^{p,k}(U))} = s \sin^2 \frac{\pi}{2s} \leq \frac{\pi^2}{4s}.$$

For the averages, by contrast, $Q_k(V^{(k)}(\varphi)) \leq Q(\varphi)$ for every k in the quadratic approximation (Proposition 89(i)).

Proof. Let $T = \text{diag}(i\pi/2, -i\pi/2)$. Put the $s/2$ links in the direction 0 at $(j, 0, 0, 0)$, j odd, $j < s$, equal to $e^{2T/s}$ and all other links equal to I . These links have offset one in the direction 0, so the tree gauge holds, and the links at even j are tree links. The path from 0 to se_0 carries e^T and every other path carries I , so one coarse link is e^T , and it lies in six coarse plaquettes, each with $\delta^2 = \|e^T - I\|_{\text{HS}}^2 = 4$: $Q_k = 24$. No fine plaquette contains two of the changed links, since the two links of a plaquette in the direction 0 differ in a transverse coordinate, so there are $3s$ nonzero fine plaquettes, each with $\delta^2 = 8 \sin^2(\pi/2s)$, and $Q = 24s \sin^2(\pi/2s)$. Gauge covariance of the path map makes the ratio gauge invariant. $\square$

A field concentrated on one thin path carries a coarse curvature of order one at a fine cost of order $1/s$. In the averages it is one transverse position among s^3 and enters with the weight s^{-3} , which is why the averages compose and the paths do not. The two properties needed for Remark 92(a) and (b) are therefore held by two different maps: the paths give a global estimate for the group-valued links on fibres that do not compose, and the averages compose on linear fibres that exist only in the quadratic approximation.

A block map with both properties cannot treat the transverse positions alike and be continuous, for any compact connected gauge group. Let $s = 2^k$ and let V be a map of the links with both ends in $B(y) \cup B(y + e_\mu)$, blocks of side s , to G . Call it *covariant* if $V(U^h) = h_1 V(U) h_2^{-1}$ for every gauge transformation h equal to h_1 on $B(y)$ and to h_2 on $B(y + e_\mu)$, and *normalised* if $V(I) = I$. The path map and its iterates are covariant, normalised and continuous. Let T be the set of the s^3 links in the direction μ from $B(y)$ to $B(y + e_\mu)$, and for $u \in G^T$ let $U(u)$ carry u_t on the link $t \in T$ and I on every other link (Fig. 13). Let $S \subset G$ be a maximal torus, of dimension ℓ ; a basis of the kernel of its exponential map identifies S with $(\mathbb{R}/2\pi\mathbb{Z})^\ell$ and its first homology with $\mathbb{Z}^\ell$.

Proposition 98. *Let $G \subset U(r)$ be a compact connected Lie group of positive dimension, let V be covariant and normalised, and let $F(u) = V(U(u))$ be continuous on G^T . Then F maps S^T into S , and the integer matrices W_t by which the maps $u_t \mapsto F(u)$, $u_t \in S$, $u_{t'} = I$ for $t' \neq t$, act on $H_1(S) = \mathbb{Z}^\ell$ satisfy*

$$\sum_{t \in T} W_t = I_\ell.$$

Hence some $W_t \neq 0$, so that the map $u_t \mapsto V(U(u))$ on fields equal to I off the link t is not homotopic to a constant on S , and, if $|T| \geq 2$, no transitive group of permutations of T leaves F invariant. For $G = SU(2)$, $\ell = 1$, the W_t are integers adding up to one, and some field equal to I off one link of T has $V = -I$.

Proof. The gauge transformation equal to g on $B(y) \cup B(y + e_\mu)$ maps $U(u)$ to $U(gug^{-1})$, since the links equal to I stay I , and the one equal to g on $B(y)$ and to I on $B(y + e_\mu)$ maps it to $U(gu)$. So $F(gug^{-1}) = gF(u)g^{-1}$ and $F(gu) = gF(u)$, and with $F(I) = I$ the second gives $F(g, \dots, g) = g$. If every u_t lies in S , every $g \in S$ fixes u under conjugation, so $F(u)$ commutes with S ; in a compact connected Lie group a maximal torus is its own centraliser [42], so $F(u) \in S$. Therefore F restricts to a continuous map $f : S^T \rightarrow S$. The first homology of S^T is $\bigoplus_{t \in T} \mathbb{Z}^\ell$, the t -th summand coming from the inclusion of the t -th factor with the identity elsewhere, so f acts on it by $(x_t)_t \mapsto \sum_t W_t x_t$. The diagonal map $S \rightarrow S^T$ acts by $x \mapsto (x, \dots, x)$, and f composed with it is the identity of S because $F(g, \dots, g) = g$. Hence $\sum_t W_t = I_\ell$. If all W_t vanished the sum would vanish, and a map of tori that is homotopic to a constant acts by zero on homology. A transitive permutation group preserving F would make all W_t equal to one matrix W , and $|T|W = I_\ell$ has no integer solution for $|T| \geq 2$. For $SU(2)$, S is a circle through $-I$, and a loop of nonzero degree in a circle is onto. $\square$

For the path map, W_t is the identity at the position of the path and zero elsewhere. Averages of the path holonomies weigh the positions alike, so by the proposition they have no continuous covariant extension to all of G^T , for $SU(N)$ and for every compact simple group. Bałaban's averages are defined only on fields whose plaquettes are small, and the large fields are split off [5, 7]. A block map with both properties for all fields is therefore discontinuous or, like the path map, not invariant under the transverse symmetries of the block. Theorem 100 below gives a continuous one whose departure from the symmetry is confined to large fields.¹⁹

The proposition leaves room for a map that treats the positions alike to first order and breaks the symmetry only where the field is large. For a coarse link (y, μ) and $x \in B(y)$ put

$$P_x = U_{x,\mu} U_{x+e_\mu,\mu},$$

the product along the straight path of length two from x to $x + 2e_\mu$. In the tree gauge it is the parallel transport from $2y$ to $2y + 2e_\mu$ through x and $x + 2e_\mu$, it transforms as a link under the gauge transformations at the block roots, and $P_{2y} = V_{y,\mu}^p$, since the first link of that path is a tree link. Let $\mathfrak{g}$ be the Lie algebra of G , $\|X\|$ the Hilbert–Schmidt norm and $\log$ the principal logarithm of unitary matrices without the eigenvalue -1 .

Lemma 99. *If g is unitary and $\delta(g, I) \leq 1$, then $\|\log g\| \leq \frac{\pi}{3} \delta(g, I)$. If X is skew-Hermitian, then $\delta(e^X, I) \leq \|X\|$.*

Proof. Let $e^{i\theta_j}$, $|\theta_j| \leq \pi$, be the eigenvalues of g . Then $\delta(g, I)^2 = \sum_j 4 \sin^2(\theta_j/2)$ and $\|\log g\|^2 = \sum_j \theta_j^2$. Each term of the first sum is at most 1, so $|\theta_j| \leq \pi/3$, and $\theta/(2 \sin(\theta/2))$ increases on $(0, \pi)$ and equals $\pi/3$ at $\theta = \pi/3$. The second statement is $|e^{i\theta} - 1| \leq |\theta|$ for the eigenvalues of X . $\square$

¹⁹ That spheres and tori carry no continuous symmetric mean goes back to Aumann [2] and Eckmann [33]; the proof above is their homology argument in gauge-covariant form, with the links of the slice T as the arguments of the mean and a maximal torus in place of the circle.

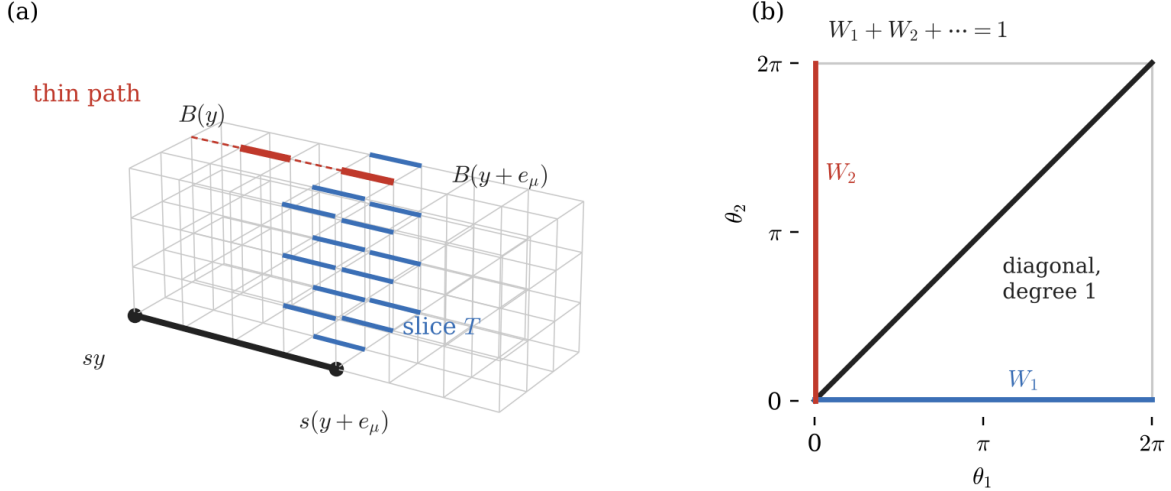


FIG. 13. (a) Two adjacent blocks $B(y)$, $B(y + e_\mu)$ of side $s = 4$, drawn in three of the four dimensions. The black segment is the straight path between the roots, the coarse link of the path map. The blue links form the slice T of Proposition 98. The red links carry the field of Proposition 97: a path whose holonomy is that of a coarse link of order one, at a fine energy of order $1/s$. (b) For $SU(2)$, two coordinates of the torus $(\mathbb{R}/2\pi\mathbb{Z})^T$ of slice fields with values in a maximal torus. The diagonal is pure gauge and is sent to a loop of degree one, so the degrees W_i of the coordinate loops add up to one.

Let $\rho_G \in (0, 1]$ be such that $\log g \in \mathfrak{g}$ for every $g \in G$ with $\delta(g, I) \leq \rho_G$.²⁰ Let $\phi(t) = \min\{1, \max\{0, 2 - 2t\}\}$, which equals 1 on $[0, \frac{1}{2}]$ and 0 on $[1, \infty)$, put

$$\chi_{y,\mu}(U) = \phi\left(\rho_G^{-1} \max_{x \in B(y)} \delta(P_x, P_{2y})\right),$$

and define the *cut-off average*

$$V_{y,\mu}^\sharp(U) = P_{2y} \exp\left(\frac{\chi_{y,\mu}(U)}{16} \sum_{x \in B(y)} \log(P_{2y}^{-1} P_x)\right), \quad (17)$$

the exponential factor being I when $\chi_{y,\mu}(U) = 0$ (Fig. 14). When every path is close to the straight one it is the average of the sixteen paths taken in the chart at P_{2y} ; when some path is far from it, it is the path map.

Theorem 100. *For every compact $G \subset U(r)$ the cut-off average is continuous, covariant and normalised, and the following hold, with $c = \pi/3$.*

- (i) *If $\delta(P_x, P_{2y}) \leq \rho_G/2$ for every $x \in B(y)$, then $V_{y,\mu}^\sharp(U) = P_{2y} \exp(\frac{1}{16} \sum_x \log(P_{2y}^{-1} P_x))$. For φ with values in $\mathfrak{g}$, vanishing on the tree links,*

$$V_{y,\mu}^\sharp(e^{t\varphi}) = I + t V_{y,\mu}(\varphi) + O(t^2),$$

with $V_{y,\mu}$ the block average of Definition 83.

- (ii) *For every field U in the tree gauge,*

$$\sum_{y,\mu} \sum_{x \in B(y)} \delta(P_x, P_{2y})^2 \leq 896 Q(U),$$

$$\sum_{y,\mu} \delta(V_{y,\mu}^\sharp(U), V_{y,\mu}^p(U))^2 \leq 56c^2 Q(U),$$

$$\sum_e \delta(U_e, R_e(V^\sharp(U)))^2 \leq (448 + 896c^2) Q(U),$$

$$Q_1(V^\sharp(U)) \leq (8 + 2688c^2) Q(U),$$

$$\text{and } 448 + 896c^2 < 1431, \quad 8 + 2688c^2 < 2956.$$

- (iii) *In the setting of Proposition 98 with $s = 2$, W_i is the identity at the position of the straight path and 0 at the seven others.*

- (iv) *If U equals I except on one link from $B(y)$ to $B(y + e_\mu)$, where it equals g with $\delta(g, I) \leq \rho_G/2$, then $V_{y,\mu}^\sharp(U) = \exp(\frac{1}{8} \log g)$, at each of the eight transverse positions of the link; the path map gives g at one position and I at the seven others.*

Proof. Write $P = P_{2y}$, $v_x = P^{-1}P_x$ and $m = \max_x \delta(v_x, I) = \max_x \delta(P_x, P)$. Where $m < \rho_G$, every $\log v_x$ lies in $\mathfrak{g}$, depends continuously on U by Lemma 99, which keeps the eigenvalues of v_x away from -1 , and the exponent of (17) lies in $\mathfrak{g}$, so $V^\sharp \in G$. Where $m \geq \rho_G$, $V^\sharp = P$. As m increases to ρ_G the exponent has norm at most $\chi c \rho_G \rightarrow 0$, so $V^\sharp$ is continuous. Let h be a gauge transformation equal to h_1 on $B(y)$ and to h_2 on $B(y + e_\mu)$. The middle point $x + e_\mu$ of each path lies in one of the two blocks, so $P_x \mapsto h_1 P_x h_2^{-1}$ and $v_x \mapsto h_2 v_x h_2^{-1}$; δ is invariant and $\log$ is equivariant under unitary conjugation, so χ is unchanged and $V^\sharp \mapsto h_1 V^\sharp h_2^{-1}$. At $U = I$ every v_x is I , so $V^\sharp(I) = I$.

²⁰ The exponential map of G is a diffeomorphism from a ball of $\mathfrak{g}$ of operator-norm radius less than π onto a neighbourhood N of I in G , on which it inverts the principal logarithm, and N contains $\{g \in G : \delta(g, I) \leq \rho\}$ for small $\rho > 0$. For $U(r)$ one may take $\rho_G = 1$. For $SU(N)$ one may take $\rho_G = \min\{1, 5/\sqrt{N}\}$: by Lemma 99 the trace of $\log g$ has modulus at most $\sqrt{N} \pi \rho_G/3 < 2\pi$, and it lies in $2\pi i\mathbb{Z}$. Some restriction is needed: on the circle $\{\text{diag}(z, z^7)\} \subset U(2)$ the point $z = e^{2\pi i/7}$ has $\delta(g, I) < 1$ but $\log g \notin \mathfrak{g}$.

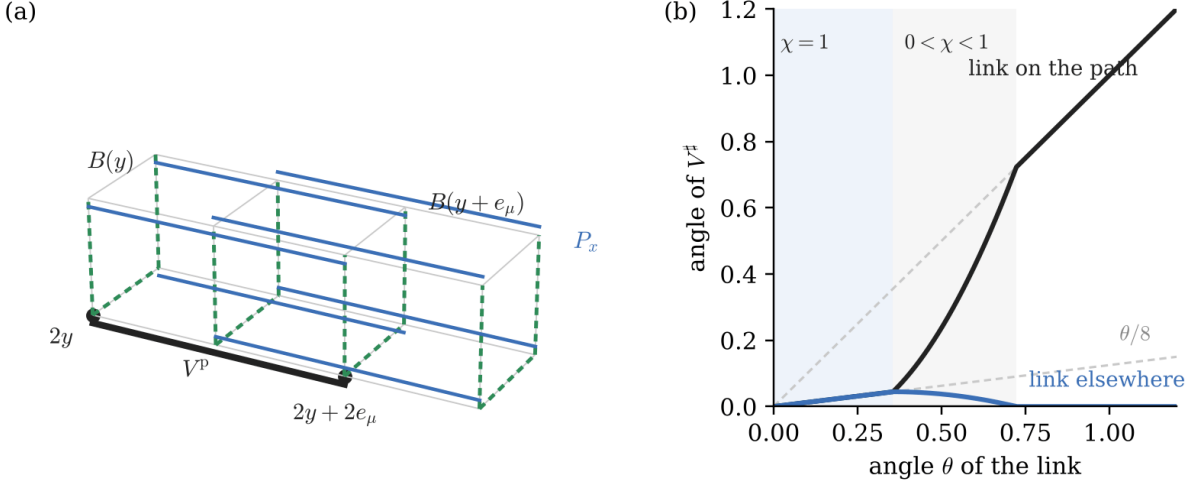


FIG. 14. (a) A block $B(y)$ of side two and the next block in the direction μ , drawn in three of the four dimensions. Green dashed: the tree links, equal to I . Blue: the paths P_x of length two from the sites of $B(y)$, each joining $2y$ to $2y + 2e_\mu$ through the trees; black: the straight path P_{2y} , the coarse link of the path map. The cut-off average is the mean of the sixteen paths in the chart at P_{2y} while every path is close to it. (b) For $SU(2)$ and one link $e^{\theta X}$, $X = \text{diag}(i, -i)$, the angle of $V^\sharp$ when the link lies on the straight path (black) or at another transverse position (blue), with $\rho_G = 1$. Both equal $\theta/8$, the block average, while $\chi = 1$; past the cut-off they are the path map, θ and 0 , so the slice windings are 1 and 0 .

(i) The first statement is $\chi = 1$. If $U = e^{t\varphi}$ with $\varphi = 0$ on the tree, then $P = I + t\varphi_{2y+e_\mu, \mu} + O(t^2)$ and $\log v_x = t(\varphi_{x, \mu} + \varphi_{x+e_\mu, \mu} - \varphi_{2y+e_\mu, \mu}) + O(t^2)$, and for small t the condition of the first statement holds. The mean over the sixteen x is $t(V_{y, \mu}(\varphi) - \varphi_{2y+e_\mu, \mu}) + O(t^2)$, since $\varphi_{2y, \mu} = 0$, and multiplying by P gives the claim.

(ii) In the tree gauge $P = v_{y, \mu} := V_{y, \mu}^p(U)$. If $o_\mu = 0$, then $U_{x+e_\mu, \mu}$ is the link leaving $B(y)$ at the transverse offset of x , and invariance of δ with the triangle inequality gives $\delta(P_x, P) \leq \delta(U_{x, \mu}, I) + \delta(U_{x+e_\mu, \mu}, v_{y, \mu})$; if $o_\mu = 1$, the roles of the two links are exchanged. Each link leaving a block in the direction μ occurs for two sites x of that block, and each link in the direction μ with $o_\mu = 0$ occurs once for the block containing it and once for the block before it. With $(a+b)^2 \leq 2a^2 + 2b^2$ and the notation of the proof of Theorem 94, the sum is at most $2(2I_0 + 2J_0) = 4D_v^2 \leq 896Q$. Next, if $\chi > 0$, Lemma 99 and the Cauchy-Schwarz inequality give

$$\begin{aligned} \delta(V^\sharp, P) &\leq \left\| \frac{\chi}{16} \sum_x \log v_x \right\| \leq \frac{c}{16} \sum_x \delta(v_x, I) \\ &\leq \frac{c}{4} \left(\sum_x \delta(P_x, P)^2 \right)^{1/2}, \end{aligned}$$

and the second bound follows from the first, $896c^2/16 = 56c^2$. The reconstruction $R(V^\sharp)$ agrees with $R(v)$ on the links with $o_\mu = 0$, and on the eight links with $o_\mu = 1$ of each block and direction $\delta(U_e, V^\sharp)^2 \leq 2\delta(U_e, v)^2 + 2\delta(V^\sharp, v)^2$; so the third sum is at most $2D_v^2 + 16 \cdot 56c^2Q \leq (448 + 896c^2)Q$. Finally, write $V^\sharp = vM$. Removing the four factors M from the holonomy of a coarse plaquette one at a time changes $\delta(\cdot, I)$ by at most $\delta(M, I)$ each time, so $\delta(V_p^\sharp, I)^2 \leq 2\delta(v_p, I)^2 + 8\sum_{\ell \in P} \delta(M_\ell, I)^2$. Each coarse link lies in six coarse plaquettes, so $Q_1(V^\sharp) \leq 2Q_1(v) + 48 \cdot 56c^2Q$, and $Q_1(v) \leq 4Q$ by Proposition 95. The numerical bounds use $c^2 < 1.0967$.

(iii) For $s = 2$ the slice T consists of the eight links leaving $B(y)$, and the links inside the blocks are I , so $P_x = u_a$ for the transverse offset a of x , two sites for each a . Let t be a position other than the straight path and $u_{t'} = I$ for $t' \neq t$. Then $P = I$, two of the v_x equal u_t and the others I , so $F(u) = \exp(\frac{1}{8}\chi \log u_t)$ with $\chi = \phi(\delta(u_t, I)/\rho_G)$. On S the maps $u_t \mapsto \exp(\frac{\lambda}{8}\chi \log u_t)$, $0 \leq \lambda \leq 1$, are continuous by the argument above and take values in S , since $\log u_t$ commutes with S and the centraliser of a maximal torus in $\mathfrak{g}$ is its Lie algebra; they join F to the constant I , so $W_t = 0$. By Proposition 98 the W_t add up to the identity, which is therefore the matrix at the straight path.

(iv) At the straight path, $P_x = g$ for the two sites with transverse offset 0 and $P_x = I$ for the others, so $V^\sharp = g \exp(-\frac{14}{16} \log g) = \exp(\frac{1}{8} \log g)$; at another position, $P = I$ and two of the P_x equal g , so $V^\sharp = \exp(\frac{2}{16} \log g)$. $\square$

The small-field part of (17) is Balaban's averaging operation [6]; the cut-off extends it to all fields with the bounds of (ii). The theorem shows how far Proposition 98 reaches. The cut-off average is continuous, covariant and coercive for every field, and it agrees with the block averages to first order. Its slice windings are those of the path map, as the proposition demands of any such map, but the departure from the transverse symmetry happens only where some path differs from the straight one by more than $\rho_G/2$. Below that the field of a single link enters with the weight $\frac{1}{8}$ of the averages, whichever position it occupies.

What the theorem does not give is composition. The constants of (ii) are those of one step, and iterating them gives a factor that grows exponentially in the number of steps; on its large-field branch the map is the path map, for which Proposition 97 rules out any bound uniform in that number. Whether the k -fold cut-off average obeys $Q_k(V^{\sharp, k}(U)) \leq CQ(U)$ and the corresponding coercivity bound with C independent of k is

open for general fields; it is the form in which Remark 92(b) remains. Section XIII C proves it, with $C = 1.001$, for fields with values in a maximal torus whose curvature stays small at every scale, and with $C = e^{12060\sigma}$ for fields of any compact group whose curvatures at the successive scales are small with sum σ .

C. Composition across scales

This subsection proves that the cut-off average composes across all scales on fields with values in a maximal torus whose curvature stays small, which includes the thin paths that defeat the path map, and then that it composes for every compact group on fields whose curvature is small, with a summable loss over the scales.

Let S be a maximal torus of $G \subset U(r)$ and $\mathfrak{s}$ its Lie algebra. The exponential maps $\mathfrak{s}$ onto S , and its kernel Λ consists of diagonalisable matrices with eigenvalues in $2\pi i\mathbb{Z}$, so every nonzero element of Λ has norm at least 2π . For an S -valued field with $\delta(U_p, I) \leq \rho_G$ on every plaquette put $\Theta_p = \log U_p$; it lies in $\mathfrak{g}$, commutes with S , and so lies in $\mathfrak{s}$. For a fine site x and a coordinate plane $\{\mu, \nu\}$ let $\square_x$ be the square of side two at x in that plane and $F_x = \sum_{p \subset \square_x} \Theta_p$, the sum of its four plaquettes, oriented as the square. Put

$$\eta_G = \rho_G/12.$$

The k -fold cut-off average is defined by putting the coarse field back into the tree gauge, by an S -valued gauge transformation, before each step. The tree gauge fixes such a transformation up to one that is constant on blocks, under which $V^\sharp$ is covariant, so the coarse plaquette holonomies $U_p^{(j)}$ of the j -th iterate $U^{(j)}$ are well defined; we write $\Theta_p^{(j)} = \log U_p^{(j)}$ when $\delta(U_p^{(j)}, I) \leq \rho_G$.

Lemma 101. *Let U be an S -valued field in the tree gauge with $\|\Theta_p\| \leq \eta_G$ for every plaquette. Then $\chi_{y,\mu}(U) = 1$ for every coarse link, and the coarse plaquette P at y in the plane $\{\mu, \nu\}$ has*

$$V^\sharp(U)_P = \exp \bar{F}_P, \quad \bar{F}_P = \frac{1}{16} \sum_{o \in \{0,1\}^4} F_{2y+o},$$

with $\|\bar{F}_P\| \leq 4 \max_p \|\Theta_p\|$, so that $\log V^\sharp(U)_P = \bar{F}_P$.

Proof. Let $\eta = \max_p \|\Theta_p\|$. Fix (y, μ) and $x = 2y + o$. The tree path τ from $2y$ to x sets the coordinates of o one at a time in increasing order, so it has at most four edges, at most three of them transverse to μ . In the tree gauge its links are I , so $P_{2y}^{-1} P_x$ is the holonomy of the loop τ , then the straight path from x to $x + 2e_\mu$, then $\tau + 2e_\mu$ backwards, then the straight path from $2y + 2e_\mu$ back to $2y$. This loop is the boundary of the surface swept by the transverse edges of τ along two steps in the direction μ : at most six distinct plaquettes, each with coefficient ± 1 . The group is abelian, so $P_{2y}^{-1} P_x = \exp d_x$ with d_x the signed sum of their Θ_p , and $\|d_x\| \leq 6\eta \leq \rho_G/2$. By Lemma 99, $\delta(P_x, P_{2y}) \leq \rho_G/2$, so $\chi = 1$, and since the eigenvalues of d_x have modulus at most $\frac{1}{2}$, $\log(P_{2y}^{-1} P_x) = d_x$.

Hence $V_{y,\mu}^\sharp = \exp(X_{y,\mu} + \bar{d}_{y,\mu})$, where $X_{y,\mu} \in \mathfrak{s}$ is any logarithm of P_{2y} and $\bar{d}_{y,\mu}$ is the mean of the d_x .

Let the sides of P be the coarse links ℓ with signs ε_ℓ , and write $d_{o,\ell}$ for the d_x of the side ℓ at the offset o . For each o the four straight paths of length two at the offset o of the four sides bound the square $\square_{2y+o}$, so

$$Y_o = \sum_\ell \varepsilon_\ell (X_\ell + d_{o,\ell}) - F_{2y+o} \in \Lambda.$$

Since $d_{o,\ell} = 0$, $\|Y_o - Y_0\| \leq 4 \cdot 6\eta + 8\eta = 32\eta \leq \frac{8}{3} < 2\pi$, so $Y_o = Y_0$ for every o . Averaging over o gives $\sum_\ell \varepsilon_\ell (X_\ell + \bar{d}_\ell) = \bar{F}_P + Y_0$, and $\exp Y_0 = I$. Finally $\|\bar{F}_P\| \leq \max_o \|F_{2y+o}\| \leq 4\eta < \pi/3$, so $\bar{F}_P$ is the principal logarithm. $\square$

Theorem 102. *Let U be an S -valued field in the tree gauge on the periodic lattice of side divisible by 2^k , with $\delta(U_p, I) \leq \rho_G$ on every plaquette, and put $\eta_0 = \max_p \|\Theta_p\|$, $A = \sum_p \|\Theta_p\|$. Suppose that*

$$\max_p \|\Theta_p^{(j)}\| \leq \eta_G \quad \text{for } 0 \leq j < k, \quad (18)$$

which holds in particular if $\eta_0 A \leq \eta_G^2$. Then the cut-off is inactive at every step, $\Theta^{(j+1)} = \bar{F}(\Theta^{(j)})$, and for $j < k$

$$\begin{aligned} \sum_p \|\Theta_p^{(j+1)}\|^2 &\leq \sum_p \|\Theta_p^{(j)}\|^2, & \sum_p \|\Theta_p^{(j+1)}\| &\leq \frac{1}{4} \sum_p \|\Theta_p^{(j)}\|, \\ \max_p \|\Theta_p^{(j+1)}\| &\leq 4 \max_p \|\Theta_p^{(j)}\|. \end{aligned}$$

Consequently

$$\begin{aligned} Q_k(V^{\sharp,k}(U)) &\leq \sum_p \|\Theta_p\|^2 \leq 1.001 Q(U), \\ Q_k(V^{\sharp,k}(U)) &\leq 16^{-k} A^2. \end{aligned}$$

Proof. The first assertions follow from Lemma 101 by induction, the S -valued gauge fixing between steps leaving the plaquette holonomies unchanged. The coarse plaquettes P in a plane together with the offsets o are in bijection with the fine squares $\square_x$ in that plane, $x = 2y + o$, and a fine plaquette lies in four such squares. With $\|\bar{F}_P\|^2 \leq \frac{1}{16} \sum_o \|F_{2y+o}\|^2$ and $\|F_x\|^2 \leq 4 \sum_{p \subset \square_x} \|\Theta_p\|^2$ this gives $\sum_p \|\bar{F}_P\|^2 \leq \frac{1}{16} \cdot 4 \cdot 4 \sum_p \|\Theta_p\|^2$; the same count without Cauchy–Schwarz gives the factor $\frac{1}{4}$ in the second inequality, and the third is $\|F_x\| \leq 4\eta$.

For the sufficient condition, let η_j and A_j be the maximum and the sum of $\|\Theta_p^{(j)}\|$. While (18) holds up to $j-1$, the inequalities give $\eta_j \leq 4^j \eta_0$ and $\eta_j \leq A_j \leq 4^{-j} A$. The product of the two bounds is $\eta_0 A \leq \eta_G^2$, so the smaller one is at most η_G , and (18) holds at j .

Finally $\delta(\exp \Theta, I) \leq \|\Theta\|$ (Lemma 99), so $Q_k(V^{\sharp,k}(U)) \leq \sum_p \|\Theta_p^{(k)}\|^2$, which is at most both $\sum_p \|\Theta_p\|^2$ and $\eta_k A_k \leq A_k^2 \leq 16^{-k} A^2$. The eigenvalues of each Θ_p are $i\theta$ with $|\theta| \leq \eta_0 \leq \eta_G \leq \frac{1}{12}$, since $\eta_0^2 \leq \eta_0 A$ in the sufficient condition and $\eta_0 \leq \eta_G$ is the case $j = 0$ of (18); and $\theta^2 \leq 1.001 \cdot 4 \sin^2(\theta/2)$ there, which gives $\sum_p \|\Theta_p\|^2 \leq 1.001 Q(U)$. $\square$

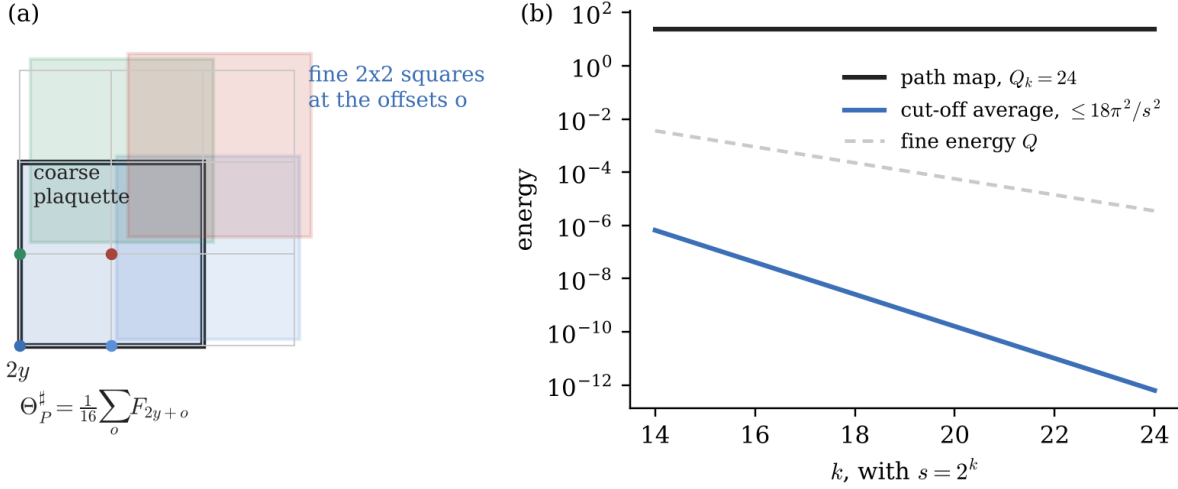


FIG. 15. (a) In the torus, the coarse plaquette of the cut-off average is the mean of the fluxes F_{2y+o} of the sixteen fine squares of side two at the offsets o (four of them drawn), whatever the gauge of the coarse links (Lemma 101). (b) For the thin-path fields of Proposition 97, the coarse energy after k steps: 24 for the path map, at most $18\pi^2/s^2$ for the cut-off average (Corollary 103); the fine energy is about $6\pi^2/s$.

Corollary 103. For $G = \text{SU}(2)$ and the fields of Proposition 97 with $s = 2^k \geq 2^{14}$,

$$Q_k(V^{\sharp,k}(U)) \leq \frac{18\pi^2}{s^2} \leq \frac{4}{s} Q(U),$$

$$Q_k(V^{\text{p},k}(U)) = 24 \geq \frac{4s}{\pi^2} Q(U).$$

Proof. Here $\rho_G = 1$ and S is the diagonal torus. The 3s nonzero plaquettes have $\Theta_p = \pm 2T/s$, of norm $\sqrt{2}\pi/s$, so $\eta_0 A = 6\pi^2/s \leq \frac{1}{144}$ when $s \geq 864\pi^2$, which holds for $s \geq 2^{14}$, and $A^2 = 18\pi^2$. Proposition 97 gives $Q(U) = 24s \sin^2(\pi/2s)$, which lies between $(6\pi^2/s)(1 - \pi^2/(12s^2))$ and $6\pi^2/s$. $\square$

So on torus-valued fields of small curvature the cut-off average behaves like the linear averages: the curvature is averaged over the sixteen offsets at every step, the energy never increases, and a thin path is spread out and contracted instead of being carried to the coarse lattice (Fig. 15).

For general fields the paths do not commute, and the logarithms carry commutators from the second order on. The next two lemmas bound these commutators by the square of the local curvature. The same composition then holds for every compact group, with a loss of at most $1 + O(\eta_j)$ at scale j .

Lemma 104. Let $Z_1, \dots, Z_n \in \mathfrak{g}$ with $r = \sum_i \|Z_i\| \leq \min\{\frac{1}{4}, \rho_G\}$, and let $g = e^{Z_1} \dots e^{Z_n}$. Then $\delta(g, I) \leq r$, $\log g \in \mathfrak{g}$, and

$$\left\| \log g - \sum_i Z_i \right\| \leq 1.6r^2.$$

Proof. The distance δ is invariant under multiplication by unitary matrices on either side, so $\delta(g, I) \leq \sum_i \delta(e^{Z_i}, I) \leq r$ by Lemma 99, and $\log g \in \mathfrak{g}$ by the choice of ρ_G . Expanding the exponentials, $g - I - \sum_i Z_i$ is a sum of products of at least two factors. Since the Hilbert–Schmidt norm is submultiplicative, its norm is at most $e^r - 1 - r \leq \frac{1}{2}r^2 e^r \leq 0.643r^2$, and $M = g - I$

has $\|M\| \leq e^r - 1 \leq 1.137r \leq 0.285$. The operator norm of M is smaller than 1. The principal logarithm is holomorphic on the disc $|z - 1| < 1$, which contains the spectrum of g , so it is given there by its power series, $\log g = \sum_{k \geq 1} (-1)^{k+1} M^k/k$. Hence

$$\|\log g - M\| \leq \frac{\|M\|^2}{2(1 - \|M\|)} \leq 0.905r^2,$$

and the two bounds add up to less than $1.6r^2$. $\square$

Let U be a G -valued field in the tree gauge with $\delta(U_p, I) \leq \rho_G$, and put $\Theta_p = \log U_p$. A change of base point conjugates Θ_p and does not change its norm. Let P be a coarse plaquette at y in the plane $\{\mu, \nu\}$. Its sides are $\ell_1 = (y, \mu)$, $\ell_2 = (y + e_\mu, \nu)$, $\ell_3 = (y + e_\nu, \mu)$ and $\ell_4 = (y, \nu)$, with initial points y_i . For an offset $o \in \{0, 1\}^4$ let $P_{i,o}$ be the path product of ℓ_i through $2y_i + o$, and put

$$K_{P,o} = P_{1,o} P_{2,o} P_{3,o}^{-1} P_{4,o}^{-1}, \quad \bar{F}_P = \frac{1}{16} \sum_o \log K_{P,o}.$$

The tree segments between consecutive sides cancel. Hence $K_{P,o}$ is the holonomy of the square $\square_{2y+o}$, carried to $2y$ along the tree. In a torus, $\log K_{P,o} = F_{2y+o}$, and $\bar{F}_P$ is the mean of Lemma 101. Let $\sigma_{P,o}$ be the sum of $\|\Theta_p\|$ over the four plaquettes of $\square_{2y+o}$.

Lemma 105. Let U be a G -valued field in the tree gauge with $\delta(U_p, I) \leq \rho_G$ and $\|\Theta_p\| \leq \eta \leq \rho_G/1000$ for every plaquette. Then $\chi_{y,\mu}(U) = 1$ for every coarse link. Every coarse plaquette has $\delta(V^\sharp(U)_P, I) \leq 0.03\rho_G$ and

$$\begin{aligned} \|\log V^\sharp(U)_P - \bar{F}_P\| &\leq 3.2r_P^2, \\ \|\bar{F}_P\| &\leq \frac{1+6.4\eta}{16} \sum_o \sigma_{P,o}, \end{aligned} \quad (19)$$

with $r_P \leq 28.26\eta$ defined in the proof. Let E, A be the sum of $\|\Theta_P\|^2$ and of $\|\Theta_P\|$, and let E', A', η' be the sum of squares, the sum and the maximum of $\|\log V^\sharp(U)_P\|$. Then

$$\sqrt{E'} \leq (1 + 6030\eta)\sqrt{E}, \quad A' \leq (1 + 9147\eta)\frac{A}{4},$$

$$\eta' \leq 4(1 + 647\eta)\eta.$$

Proof. Paths. Fix a coarse link (y, μ) and $x = 2y + o$. As in the proof of Lemma 101, $P_{2y}^{-1}P_x$ is the holonomy of a loop bounding a surface of at most six plaquettes, swept by the transverse edges of the tree path. Build the surface one plaquette at a time. This writes the holonomy as a product of at most six factors $hU_p^{\pm 1}h^{-1} = \exp(\pm h\Theta_p h^{-1})$. Let $\sigma_{(y, \mu), o} \leq 6\eta$ be the sum of their $\|\Theta_p\|$. By Lemma 104, $\delta(P_x, P_{2y}) \leq 6\eta \leq \rho_G/2$, so $\chi = 1$. The logarithm $D_{(y, \mu), o} = \log(P_{2y}^{-1}P_x)$ lies in $\mathfrak{g}$, with

$$\|D_{(y, \mu), o}\| \leq (1 + 9.6\eta)\sigma_{(y, \mu), o}.$$

In the same way $K_{P, o}$ is a product of four such factors, so $\|\log K_{P, o}\| \leq (1 + 6.4\eta)\sigma_{P, o}$. Averaging over o gives the second inequality of (19).

The plaquette. Write $D_{i, o}$ for $D_{\ell_i, o}$. Then $P_{i, o} = P_{i, 0}e^{D_{i, o}}$ and $D_{i, 0} = 0$, and $V_{\ell_i}^\sharp = P_{i, 0}e^{m_i}$ with $m_i = \frac{1}{16}\sum_o D_{i, o}$. The paths and $V^\sharp$ transform as links under gauge transformations constant on blocks. These transformations preserve the tree gauge and conjugate $D_{i, o}$, m_i , $K_{P, o}$ and $V_P^\sharp$ by fixed elements. Choose one with $P_{1, 0} = P_{2, 0} = P_{3, 0} = I$. Then $P_{4, 0} = K_{P, 0}^{-1}$, and with $c = \log K_{P, 0}$,

$$V_P^\sharp = e^{m_1}e^{m_2}e^{-m_3}e^{-m_4}e^c,$$

$$K_{P, o} = e^{D_{1, o}}e^{D_{2, o}}e^{-D_{3, o}}e^{-D_{4, o}}e^c.$$

Put $r_P = \sum_i \max_o \|D_{i, o}\| + \|c\|$. Since $\|m_i\| \leq \max_o \|D_{i, o}\|$, the five exponents in each product have norms adding up to at most r_P , and

$$r_P \leq 24(1 + 9.6\eta)\eta + 4(1 + 6.4\eta)\eta \leq 28.26\eta.$$

Lemma 104 gives $\delta(V_P^\sharp, I) \leq r_P \leq 0.03\rho_G$ and

$$\log V_P^\sharp = m_1 + m_2 - m_3 - m_4 + c + R,$$

$$\log K_{P, o} = D_{1, o} + D_{2, o} - D_{3, o} - D_{4, o} + c + R_o,$$

with $\|R\|, \|R_o\| \leq 1.6r_P^2$. Average the second line over o and subtract it from the first; since m_i is the mean of the $D_{i, o}$, this is the first inequality of (19).

Sums. Each fine plaquette lies in four squares $\square_x$, and the pairs (P, o) are in bijection with these squares. The second inequality of (19) and Cauchy–Schwarz therefore give $\sum_P \|\bar{F}_P\|^2 \leq (1 + 6.4\eta)^2 E$; without Cauchy–Schwarz they give $\sum_P \|\bar{F}_P\| \leq (1 + 6.4\eta)A/4$. For the remainders, let $\mathcal{U}_\ell$ be the set of plaquettes of the surfaces of the side $\ell = (y, \mu)$ for all sixteen offsets. These lie in planes $\{\mu, \alpha\}$, $\alpha \neq \mu$, at μ -positions $2y_\mu$, $2y_\mu + 1$, $2y_\mu + 2$, and there are at most 36 of them. A plaquette lies in $\mathcal{U}_\ell$ for at most four coarse links. With $m'_\ell = \sum_{P \in \mathcal{U}_\ell} \|\Theta_P\|$,

$$r_P \leq 1.0096 \left(\sum_{\ell \subset P} m'_\ell + \sigma_{P, 0} \right).$$

A coarse link is a side of six coarse plaquettes, and a plaquette lies in $\square_{2y}$ for at most one coarse plaquette. Therefore

$$\sum_P r_P \leq 1.0096(6 \cdot 4A + A) \leq 25.24A,$$

$$\sum_P r_P^2 \leq 1.0096^2 \cdot 5(6 \cdot 36 \cdot 4E + 4E) \leq 4424E.$$

Hence, with $r_P \leq 28.26\eta$,

$$\sum_P 3.2r_P^2 \leq 3.2 \cdot 28.26 \cdot 25.24\eta A \leq 2283\eta A,$$

$$\left(\sum_P (3.2r_P^2) \right)^{1/2} \leq 3.2 \cdot 28.26\eta (4424E)^{1/2} \leq 6015\eta\sqrt{E}.$$

The triangle inequality, for the Euclidean norm over P in the first case, gives $\sqrt{E'} \leq (1 + 6.4\eta + 6015\eta)\sqrt{E}$ and $A' \leq (1 + 6.4\eta)A/4 + 2283\eta A$. For the maximum, $\eta' \leq (1 + 6.4\eta)4\eta + 3.2(28.26\eta)^2 \leq 4(1 + 647\eta)\eta$. $\square$

For a G -valued field, the k -fold cut-off average is defined as before, with G -valued gauge transformations. The tree gauge fixes them up to transformations constant on blocks, under which $V^\sharp$ is covariant. These transformations conjugate the plaquette holonomies and leave the norms $\|\Theta_P^{(j)}\|$ unchanged.

Theorem 106. *Let $G \subset \mathbf{U}(r)$ be compact and let U be a G -valued field on the periodic lattice of side divisible by 2^k , with $\delta(U_P, I) \leq \rho_G$ on every plaquette. Let η_j, A_j, E_j be the maximum, the sum and the sum of squares of $\|\Theta_P^{(j)}\|$ over the plaquettes of the j -th iterate.*

(i) *If $\eta_j \leq \rho_G/1000$ for $0 \leq j < k$, the cut-off is inactive at every step. With $\sigma = \sum_{j < k} \eta_j$,*

$$E_k \leq e^{12060\sigma}E_0, \quad A_k \leq e^{9147\sigma}4^{-k}A_0,$$

$$\eta_k \leq e^{647\sigma}4^k\eta_0.$$

(ii) *If $\eta_0 A_0 \leq 2 \cdot 10^{-10} \rho_G^2$, then (i) applies for every k with $\sigma \leq 7.25\sqrt{\eta_0 A_0}$, and*

$$Q_k(V^{\sharp, k}(U)) \leq e^{8.8 \cdot 10^4 \sqrt{\eta_0 A_0}} E_0 \leq 1.0001 e^{8.8 \cdot 10^4 \sqrt{\eta_0 A_0}} Q(U),$$

$$Q_k(V^{\sharp, k}(U)) \leq e^2 16^{-k} A_0^2.$$

Proof. (i) Apply Lemma 105 at each step to the field in the tree gauge; the gauge fixing does not change the norms. The lemma gives $\delta(U_P^{(j+1)}, I) \leq \rho_G$, so $\Theta^{(j+1)}$ is defined, and with $1 + t \leq e^t$ the three inequalities of the lemma multiply over the steps.

(ii) Suppose that for $i < j$ we have $\eta_i \leq \rho_G/1000$ and $\sum_{i < j} \eta_i \leq 7.25\sqrt{\eta_0 A_0} \leq 1/9747$. Then (i) gives $A_j \leq e 4^{-j} A_0$ and $\eta_j \leq e 4^j \eta_0$, and $\eta_j \leq A_j$. So $\eta_j \leq e \min\{4^j \eta_0, 4^{-j} A_0\} \leq e \sqrt{\eta_0 A_0} \leq \rho_G/1000$. Let i_* be the last index with $4^{i_*} \eta_0 \leq 4^{-i_*} A_0$. The terms of $\min\{4^i \eta_0, 4^{-i} A_0\}$ increase by the factor 4 up to i_* and decrease by the factor 4 after it. The terms at i_* and $i_* + 1$ are at most $\sqrt{\eta_0 A_0}$, so the sum over all i is at most $\frac{8}{3}\sqrt{\eta_0 A_0}$, and $\sum_{i \leq j} \eta_i \leq \frac{8e}{3}\sqrt{\eta_0 A_0} \leq 7.25\sqrt{\eta_0 A_0}$.

This completes the induction. Then $E_k \leq e^{12060 \cdot 7.25 \sqrt{\eta_0 A_0}} E_0$. By Lemma 99, $Q_k(V^{\sharp,k}(U)) \leq E_k$, and $E_k \leq \eta_k A_k \leq A_k^2 \leq e^2 16^{-k} A_0^2$. Finally the eigenvalues $i\theta$ of each Θ_p have $|\theta| \leq \eta_0 \leq \sqrt{\eta_0 A_0} < 10^{-4}$, and $\theta^2 \leq 1.0001 \cdot 4 \sin^2(\theta/2)$ there. $\square$

For a torus Theorem 102 is sharper, and its hypothesis is weaker; the content of Theorem 106 is the non-abelian case. Neither theorem gives the supremum over boundary conditions in (D). Theorem 100 has the constants of one step. Theorems 102 and 106 need the curvature to be small at every scale: in Theorem 106 the losses $e^{6030\eta_j}$ must have a bounded product, and a field whose curvature is smooth on a scale ℓ grows by a factor four per step until the coarse spacing reaches ℓ . The large fields, where the cut-off is active, and the Wilson measure itself are outside both; Sec. XII D shows where the large fields lie under that measure.

D. Large fields under the Wilson measure

Theorem 106 needs every plaquette to be small. This subsection shows that, under the $SU(2)$ Wilson measure, the plaquettes that are not small form clusters of microscopic size, and that the cut-off average departs from the average only next to them. It also shows that such plaquettes are present along the one-loop tuning, so the composition has to pass through them.

Let $\Lambda_L = (\mathbb{Z}/L\mathbb{Z})^4$ with $L = 2^m \geq 4$, and let μ_L be the $SU(2)$ Wilson measure on Λ_L , with density proportional to $e^{-\beta_W S(U)}$ with respect to product Haar measure, where

$$S(U) = \sum_p s_p, \quad s_p = 1 - \frac{1}{2} \text{Tr} U_p = \frac{1}{4} \delta(U_p, I)^2.$$

This is the Wilson measure of Sec. III on the torus, without gauge fixing; it has $4L^4$ links and $6L^4$ plaquettes. For $0 < \rho \leq 1$ call a plaquette *bad* if $\delta(U_p, I) > \rho$. The closed unit cube $\mathcal{C}_x = x + [0, 1]^4$ contains 24 plaquettes. Let B_x be the event that one of them is bad. A plaquette lies in exactly four unit cubes, one for each choice of the two transverse coordinates of the cube's corner.

Lemma 107. *Let Z_L be the normalising constant of μ_L and $C_0 = \frac{1}{2} + \frac{2}{3} \log(24\pi^3)$. For $\beta_W \geq 1$,*

$$-\log Z_L \leq 6L^4 (\log \beta_W + C_0), \quad \mathbb{E}_{\mu_L} S \leq 6L^4 \frac{\log \beta_W + C_0}{\beta_W}.$$

Proof. Write $U = \cos \theta I + \sin \theta J$ with $0 \leq \theta \leq \pi$ and $J^2 = -I$. The angle θ of $g^{-1}h$ is a bi-invariant distance, and under Haar measure θ has density $\frac{2}{\pi} \sin^2 \theta$. Put $r = (4\sqrt{\beta_W})^{-1}$. If every link has angle at most r , every plaquette has angle at most $4r \leq 1$ and $s_p \leq (4r)^2/2 = 1/(2\beta_W)$, so $S \leq 3L^4$. Each link angle is at most r with probability

$$\frac{2}{\pi} \int_0^r \sin^2 t \, dt \geq \frac{8r^3}{3\pi^3} = \frac{1}{24\pi^3 \beta_W^{3/2}},$$

by $\sin t \geq 2t/\pi$. Hence $Z_L \geq e^{-3L^4} (24\pi^3 \beta_W^{3/2})^{-4L^4}$, which is the first bound. The function $F(b) = -\log Z_L$ at the coupling b

is concave with $F(0) = 0$ and $F'(b) = \mathbb{E} S$, so $bF'(b) \leq F(b)$, which is the second. $\square$

Theorem 108. *For $\beta_W \geq 1$, $0 < \rho \leq 1$ and every set A of unit cubes,*

$$\mu_L \left(\bigcap_{x \in A} B_x \right) \leq q^{|A|},$$

$$q = q_\rho(\beta_W) = \min\{1, \exp(6(\log \beta_W + C_0) - \beta_W \rho^2/16)\}.$$

Proof. The measure μ_L is reflection positive in every hyperplane through sites [64]: the plaquettes are either in the hyperplane or on one side of it, so conditionally on the links in the hyperplane the weight factors into two mirror images, and reflection preserves Haar measure and $\text{Tr} U_p$. On a torus of side 2^m this gives the chessboard estimate for functions of the links of closed unit cubes [35], [20, Section 5.3]:

$$\mu_L \left(\bigcap_{x \in A} B_x \right) \leq \mu_L \left(\bigcap_{x \in \Lambda_L} B_x \right)^{|A|/L^4},$$

since the reflection of B_x into another cube is the event B_y of that cube. If every cube is bad, there are at least $L^4/4$ bad plaquettes, so $S \geq \rho^2 L^4/16$ and the event has unnormalised weight at most $e^{-\beta_W \rho^2 L^4/16}$. Dividing by Z_L and using Lemma 107 gives the bound. $\square$

Corollary 109. *Join two unit cubes when their corners are at ℓ^∞ distance one on the torus. For an integer $r \geq 1$,*

$$\mu_L \{\text{some connected set of } r \text{ bad cubes}\} \leq L^4 (6400q)^r.$$

Hence, on a torus of fixed physical side La , if $a \rightarrow 0$ along $L = 2^m$ and $\beta_W(a) \rightarrow \infty$, the physical diameters of all connected clusters of bad cubes tend to 0 in probability, for every fixed ρ .

Proof. A connected set of r cubes containing a given cube is recovered from a walk of length $2(r-1)$ around a spanning tree, and each cube has 80 neighbours, so there are at most $80^{2(r-1)} \leq 6400^r$ such sets. Apply Theorem 108 and sum over the L^4 roots. A cluster of physical diameter at least R contains a connected set of $\lceil R/a \rceil$ cubes, and once $6400q \leq \frac{1}{2}$ the bound is at most $(La/a)^4 2^{-R/a} \rightarrow 0$. $\square$

The bad cubes control where the cut-off average differs from the average. For a coarse link (y, μ) let $\mathcal{Z}_{y, \mu}$ be the 24 unit cubes with corners $2y + te_\mu + w$, $t \in \{0, 1, 2\}$, $w \in \{0, 1\}^4$, $w_\mu = 0$ (Fig. 16).

Corollary 110. *Let $\rho = \frac{1}{12}$. If no cube of $\mathcal{Z}_{y, \mu}$ is bad, then $\chi_{y, \mu}(U) = 1$ in the tree gauge, and $V_{y, \mu}^\sharp$ is the average of the sixteen paths in the chart at P_{2y} . A bad cube lies in $\mathcal{Z}_{y, \mu}$ for at most eight coarse links, so the expected fraction of coarse links with $\chi < 1$ is at most $4.5 \cdot 10^5 (\log \beta_W + C_0)/\beta_W$. Along the limits of Corollary 109, with probability tending to one these links lie within physical distance $2a$ of clusters of bad cubes of physical diameter less than any fixed $R > 0$.*

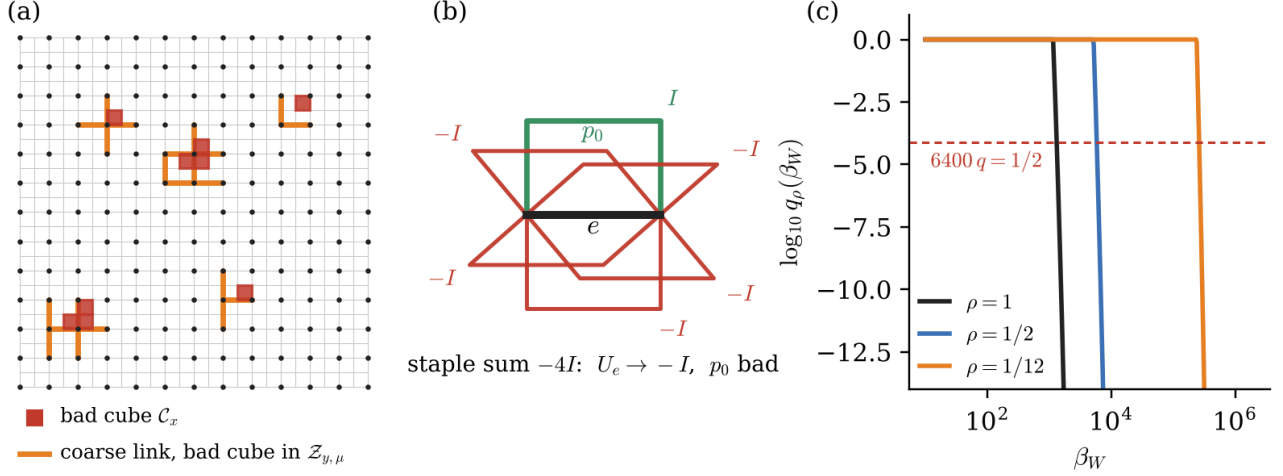


FIG. 16. Large fields under the SU(2) Wilson measure. (a) A plane section of the lattice: bad unit cubes (red) form clusters, and the cut-off average can differ from the path average only on coarse links (orange) whose 24 cubes $\mathcal{Z}_{y,\mu}$ meet a bad cube (Corollary 110). (b) The frustrated exterior of a link e : the staple of p_0 is I and the other five staples are $-I$, so U_e is driven to $-I$ and p_0 is bad. (c) The weight $q_\rho(\beta_W)$ of Theorem 108 for three values of ρ ; below the dashed line the cluster bound of Corollary 109 is summable.

Proof. As in the proof of Lemma 105, $P_{2y}^{-1}P_x$ is a product of at most six conjugates of plaquette holonomies or their inverses. These plaquettes lie in planes $\{\mu, \alpha\}$ with corners $2y + te_\mu + w$ as above, and a plaquette lies in the unit cube at its corner. So $\delta(P_x, P_{2y}) \leq 6\rho = \frac{1}{2}$, and $\chi = 1$. A cube with corner z lies in $\mathcal{Z}_{y,\mu}$ only if $z_\mu - 2y_\mu \in \{0, 1, 2\}$ and $z_\alpha - 2y_\alpha \in \{0, 1\}$ for $\alpha \neq \mu$, which leaves at most two coarse links in each direction. There are $L^4/4$ coarse links, the number of bad cubes is at most four times the number of bad plaquettes, and a bad plaquette has $s_p > \rho^2/4$, so by Lemma 107 the expected fraction is at most $4 \cdot 8 \cdot 4 \cdot 24 \cdot 144 (\log \beta_W + C_0)/\beta_W$. The last assertion is Corollary 109. $\square$

Small plaquettes everywhere, which Theorem 106 assumes, are nevertheless improbable along the one-loop tuning. For SU(2), $\beta_W = 4g^{-2}$, and the one-loop coupling of Remark 92(c) at spacing $a = 1/L$ gives $\beta_W = 8b_0 \log L + O(1) = \frac{11}{3\pi^2} \log L + O(1)$.

Proposition 111. *Let $0 < \rho \leq 1$, $v_* = \frac{2}{\pi} \int_{\arccos(3/4)}^{\pi/3} \sin^2 t \, dt$ and $p = v_* e^{-3\beta_W}$. Conditionally on all other links, a plaquette through a link e is bad with probability at least p . Hence*

$$\mu_L\{\text{no plaquette is bad}\} \leq e^{-\lfloor L/3 \rfloor^4 p},$$

which tends to 0 when $\beta_W \leq B \log L + O(1)$ with $B < 4/3$, in particular along the one-loop tuning.

Proof. Identify SU(2) with the unit quaternions. Conditionally on the other links, U_e has density proportional to $e^{\kappa h \cdot u}$ with $0 \leq \kappa \leq 6\beta_W$ and $|h| = 1$, the six staples contributing at most β_W each. The plaquette holonomy is U_e or its inverse multiplied by a fixed element, so it has a density of the same form. Writing the density as a superposition of indicators of caps centred at h , the probability of a cap about I is largest when $h = I$, and for $h = I$ it increases with κ , because its derivative is

the covariance of an increasing function of $\text{Tr } U$ with $\frac{1}{2} \text{Tr } U$. So the probability of being bad is at least that of the angle exceeding $\arccos(1 - \rho^2/4) \leq \arccos(3/4)$ under the density proportional to $e^{6\beta_W \cos t} \sin^2 t$, which is at least p : bound the numerator from below on $[\arccos(3/4), \pi/3]$, where $\cos t \geq \frac{1}{2}$, and the denominator by $e^{6\beta_W} \pi/2$.

Let $M = \lfloor L/3 \rfloor$. Choose the links in direction 0 at the sites $3z$, $z \in \{0, \dots, M-1\}^4$, and their plaquettes in the plane $\{0, 1\}$. No chosen plaquette contains another chosen link. Removing one chosen link at a time therefore bounds the probability that all chosen plaquettes are good by $(1-p)^{M^4}$. With $\beta_W = B \log L + O(1)$ the exponent is of order L^{4-3B} . $\square$

The joint bound of Theorem 108 is not a conditional one, and (D) needs a supremum over boundary conditions. The two differ genuinely. Fix a link e and a plaquette p_0 through it. Put all links on the six staples of e at I , except the middle links of the five staples other than that of p_0 , which are $-I$. Then the conditional density of U_e is proportional to $e^{-2\beta_W \text{Tr } U_e}$, and p_0 , whose holonomy is $U_e^{\pm 1}$, is bad with conditional probability tending to 1 as $\beta_W \rightarrow \infty$, for every $\rho < 2\sqrt{2}$. Nearby exteriors behave the same way. So the clusters of Corollary 109 are small in probability, but the boundary response of a cube to a frustrated exterior is not small, and the composition near the bad clusters and the conditional estimate in (D) remain open.

ACKNOWLEDGEMENTS

This paper rests on the work of many earlier scientists. Lattice gauge theory is due to Wilson [77] and its Hamiltonian form to Kogut and Susskind [51]; reflection positivity and the reconstruction of the transfer matrix to Osterwalder and Schrader [62, 63], and their lattice form, with the first proof of a gap at strong coupling, to Osterwalder and Seiler [64, 68].

The uniqueness method is Dobrushin's [29–31], the curvature method is that of Bakry and Émery [3, 4] in the form of Shen, Zhu and Zhu [69], the gauge fixing on a forest is Creutz's [26], and the curvature of Lie groups is Milnor's [60]; the localisation estimate of Sec. XII A is that of Combes and Thomas [27]. The analysis of the single-link law uses results of Barta [9], Biernacki and Krzyż [19], McShane [59], Stein and Weiss [72] and Watson [76], and the continuum problem is the one formulated by Jaffe and Witten [47] and studied by Bałaban [5–7]. The Yang–Mills manuscripts of Sussman and collaborators [73, 74] were read in the preparation of Sec. IX.

DATA AND CODE AVAILABILITY

The proofs are analytic except for the evaluations in Propositions 34 and 35 and Lemma 38 and the numerical bounds of Lemma 16, which are computations in interval arithmetic with rigorous error bounds. The programs of these steps, in Python with python-flint (FLINT and Arb) and a C kernel for products of Taylor polynomials, their outputs and the list of certified boxes are in the repository <https://github.com/creelie/4Y-M>, release v4.8.0, together with a Lean 4 certificate of the line field and independent exact or numerical checks of the other finite statements.

The computations, the checks and parts of the drafting were done with AI assistance (Claude Opus 5.5, Claude Fable 5.1, GPT-6.1 Sol Ultra and GPT-6 Astra Ultra) under the direction of Deep Bhattacharjee. The combination of the forest gauge with the sharp quarter behind Theorem 1 was first assembled by Anaxia, the reasoning engine of the BLLM system [12, 13]; the line field, the free-boundary limit, the ceiling, the curvature bound, the SU(3) part, the plaquette blocks and the certified constants are the authors' own. All three authors approved the final text.

DECLARATIONS

Funding. None. *Competing interests.* None. *Ethics approval.* Not applicable: the work involves no human participants, animals or personal data.

Appendix A: The line field

Table III lists, for each residue class of x modulo two, the direction $\sigma(x)$ and the permutation $\pi_x(\mu) = \sigma(x + e_\mu)$. In every row π_x fixes $\sigma(x)$ and is a three-cycle on the other directions, which is the condition of Lemma 8.

Appendix B: The comparison lemma

Let (G, δ) be a compact group with a bi-invariant metric of diameter D , $\Omega = G^A$, and $C_{ee'}$ the coefficients of Sec. III, with row and column sums at most $\alpha < 1$ and $C_{ee'} = 0$ unless e, e'

TABLE III. The line field: x modulo two (digits $x_0x_1x_2x_3$), the direction $\sigma(x)$ and the values $\sigma(x + e_\mu)$ for $\mu = 0, 1, 2, 3$.

x	$\sigma(x)$	$\mu = 0$	$\mu = 1$	$\mu = 2$	$\mu = 3$
0000	0	0	2	3	1
1000	0	0	3	1	2
0100	2	3	0	2	1
1100	3	2	0	1	3
0010	3	1	2	0	3
1010	1	3	1	0	2
0110	2	1	3	2	0
1110	1	2	1	3	0
0001	1	2	1	3	0
1001	2	1	3	2	0
0101	1	3	1	0	2
1101	3	1	2	0	3
0011	3	2	0	1	3
1011	2	3	0	2	1
0111	0	0	3	1	2
1111	0	0	2	3	1

share a plaquette, so that $|e - e'| \leq 1$ for the distance of link midpoints. Fix $\lambda > 0$ with $\theta = \alpha e^\lambda < 1$.

Lemma 112. *Let $S \subset A$, possibly infinite, and let μ_1, μ_2 be probability measures on Ω invariant under γ_e^F for every $e \notin S$. Let $T \subset A \setminus S$ be finite, $w_b = e^{-\lambda \text{dist}(b, T)}$, and suppose $\sum_{z \in S} w_z < \infty$. There is a coupling of μ_1, μ_2 whose link distances $\rho(b) = \mathbb{E} \delta(\sigma_b, \sigma'_b)$ satisfy*

$$\sum_{b \notin S} w_b \rho(b) \leq \frac{\theta D}{1 - \theta} \sum_{z \in S} w_z.$$

If $S = \emptyset$ then $\mu_1 = \mu_2$.

Proof. The couplings form a weakly compact set and $\Psi(\pi) = \sum_b w_b \mathbb{E}_\pi \delta(\sigma_b, \sigma'_b)$ is weakly continuous because $\sum_b w_b < \infty$; let π^* be a minimiser. For $e \notin S$ replace the pair (σ_e, σ'_e) by a sample of an optimal coupling of $\gamma_e^F(\cdot | \sigma)$ and $\gamma_e^F(\cdot | \sigma')$, chosen measurably [75, Cor. 5.22]. Both μ_i are invariant, so the result is again a coupling, and changing the finitely many links that enter γ_e^F one at a time gives $W_1 \leq \sum_{e'} C_{ee'} \delta(\sigma_{e'}, \sigma'_{e'})$. Minimality gives $\rho(e) \leq \sum_{e'} C_{ee'} \rho(e')$. Multiply by w_e and sum over $e \notin S$; where $C_{ee'} \neq 0$, $w_e \leq e^\lambda w_{e'}$, so

$$\sum_{e \notin S} w_e \rho(e) \leq \theta \left[\sum_{e' \notin S} w_{e'} \rho(e') + \sum_{z \in S} w_z \rho(z) \right],$$

all sums are finite, and $\rho \leq D$ on S . If $S = \emptyset$, $\rho \equiv 0$. $\square$

Corollary 113. *Let μ be the Gibbs measure of γ^F , F a local function with support S_F , L_F -Lipschitz in each link, and H bounded with finite support S_H disjoint from S_F . Then*

$$|\text{Cov}_\mu(F, H)| \leq \|H\|_\infty L_F \frac{\theta D}{1 - \theta} |S_H| e^{-\lambda \text{dist}(S_F, S_H)}.$$

Proof. Condition on the links of S_H [50, Ch. 3]; the conditional measures $\mu^\sigma = \mu(\cdot \mid \sigma_{S_H})$ are invariant under γ_e^F for $e \notin S_H$. Lemma 112 with $S = S_H$ and $T = S_F$, for which $w_b = 1$ on S_F and $w_z \leq e^{-\lambda \text{dist}(S_F, S_H)}$ on S_H , gives $|\mu^\sigma(F) - \mu^{\sigma'}(F)| \leq L_F \sum_{b \in S_F} \rho(b) \leq L_F \frac{\theta D}{1-\theta} |S_H| e^{-\lambda \text{dist}(S_F, S_H)}$ for any two boundary values. Since H depends only on σ_{S_H} , $\text{Cov}_\mu(F, H) = \iint H(\sigma) (\mu^\sigma(F) - \mu^{\sigma'}(F)) d\mu(\sigma) d\mu(\sigma')$, and the bound follows. $\square$

Lemma 114. *Let the blocks partition A , let C_{Bj} be the block coefficients of Sec. V, with $\sum_B C_{Bj} \leq \alpha < 1$ for every j and $C_{Bj} = 0$ unless every link of B lies within distance 2 of j , and fix $\lambda > 0$ with $\theta = \alpha e^{2\lambda} < 1$. Let $S \subset A$ be a union of blocks and let μ_1, μ_2 be invariant under γ_B^θ for every block $B \not\subset S$. With T, w_b as in Lemma 112 and $\sum_{z \in S} w_z < \infty$, there is a coupling with $\sum_{b \notin S} w_b \rho(b) \leq \frac{\theta D}{1-\theta} \sum_{z \in S} w_z$, and $\mu_1 = \mu_2$ if $S = \emptyset$.*

Proof. Take a minimiser π^* of Ψ as before. For a block $B \not\subset S$ replace (σ_B, σ'_B) by an optimal coupling, for the metric d , of the two block laws; it is again a coupling, and changing the boundary links one at a time bounds the expected new d -distance by $\sum_j C_{Bj} \rho(j)$. Put $\bar{w}_B = \max_{i \in B} w_i$. Minimality gives

$$\sum_{i \in B} w_i \rho(i) \leq \bar{w}_B \sum_j C_{Bj} \rho(j),$$

and $\bar{w}_B \leq e^{2\lambda} w_j$ whenever $C_{Bj} \neq 0$. Summing over the blocks outside S and using the column sums,

$$\sum_{b \notin S} w_b \rho(b) \leq \theta \left[\sum_{j \notin S} w_j \rho(j) + \sum_{z \in S} w_z \rho(z) \right],$$

which gives the claim as in Lemma 112. $\square$

Lemma 115. *Let $n \geq 2$ be even, let $V_x, x \in 2\mathbb{Z}^4$, be the translates of Sec. X, and suppose that $D(n, \alpha)$ holds.*

- (a) *If a probability measure μ on G^A is invariant under γ_e^F for every e in a finite set V , it is invariant under γ_V .*
- (b) *Fix $\lambda > 0$ with $\theta = \alpha e^{\lambda \rho_n} < 1$. Let $S \subset A$, and let μ_1, μ_2 be invariant under γ_x for every x with $V_x \cap S = \emptyset$. With T, w_b as in Lemma 112 and $\sum_{z \in S} w_z < \infty$, there is a coupling with*

$$\sum_{b \notin S} w_b \rho(b) \leq \frac{D(1 + c_n e^{\lambda \rho_n})}{1 - \theta} \sum_{z \in S} w_z,$$

and $\mu_1 = \mu_2$ if $S = \emptyset$.

Proof. (a) Let ν_η be the conditional law of Q_V given the links η outside V . For bounded f on G^V and bounded g of η , $\gamma_e^F(fg) = g \gamma_e^F f$ for $e \in V$, so invariance gives $\int g(\eta) (\nu_\eta(\gamma_e^F f) - \nu_\eta(f)) d\mu = 0$ for all g ; taking f in a countable dense set of continuous functions, ν_η is invariant under every $\gamma_e^F(\cdot \mid \cdot, \eta)$, $e \in V$, for μ -almost every η . So is $\gamma_V(\cdot \mid \eta)$, whose conditional laws of single links are the γ_e^F . The composition P_η of the kernels γ_e^F , $e \in V$, in a fixed order has a density with respect to Haar measure on G^V bounded below by

a positive constant, because each single-link density is at least $e^{-12N^2\beta}$ times its maximum over the group. A Markov kernel with such a lower bound has at most one invariant probability measure (Doebelin's condition), so $\nu_\eta = \gamma_V(\cdot \mid \eta)$.

(b) Let π^* minimise Ψ as in Lemma 112 and call x admissible if $V_x \cap S = \emptyset$. For admissible x replace $(\sigma_{V_x}, \sigma'_{V_x})$ by an optimal coupling, for d_{V_x} , of the two block laws; both μ_i are invariant, so the result is a coupling, and changing the boundary links one at a time bounds the new expected d_{V_x} -distance by $\sum_j C_{V_x, j} \rho(j)$. With $\bar{w}_x = \max_{i \in V_x} w_i$, minimality gives

$$\sum_{i \in V_x} w_i \rho(i) \leq \bar{w}_x \sum_{j \notin V_x} C_{V_x, j} \rho(j),$$

and $\bar{w}_x \leq e^{\lambda \rho_n} w_j$ whenever $C_{V_x, j} \neq 0$. Let S' be the set of links outside S that lie in a translate meeting S ; each is within distance ρ_n of a link of S , so $\sum_{i \in S'} w_i \leq c_n e^{\lambda \rho_n} \sum_{z \in S} w_z$. A link outside $S \cup S'$ lies in N_n translates, all admissible. Summing over admissible x and using $\sum_{x: j \notin V_x} C_{V_x, j} = K_c(j) \leq \alpha N_n$,

$$N_n \sum_{i \notin S \cup S'} w_i \rho(i) \leq \theta N_n \left[\sum_{i \notin S \cup S'} w_i \rho(i) + \sum_{i \in S \cup S'} w_i \rho(i) \right].$$

All sums are finite and $\rho \leq D$, so $\sum_{i \notin S \cup S'} w_i \rho(i) \leq \frac{\theta D}{1-\theta} (1 + c_n e^{\lambda \rho_n}) \sum_{z \in S} w_z$; adding $\sum_{i \in S'} w_i \rho(i) \leq D c_n e^{\lambda \rho_n} \sum_{z \in S} w_z$ gives the bound. If $S = \emptyset$ then $S' = \emptyset$ and $\rho \equiv 0$. $\square$

The covariance bound of Corollary 113 follows from part (b) in the same way, with $\theta D/(1-\theta)$ replaced by $D(1 + c_n e^{\lambda \rho_n})/(1-\theta)$, since the conditional measures μ^σ there are invariant under γ_x for $V_x \cap S_H = \emptyset$ by part (a).

Appendix C: The plaquette blocks

Table IV lists the block plaquettes of Sec. V modulo $2\mathbb{Z}^4$: the plaquette at x in the plane $\{\mu, \nu\}$ is a block plaquette exactly when x modulo two and $\{\mu, \nu\}$ appear in the table. Eight lie in the plane $\{0, 1\}$ and eight in $\{2, 3\}$; together they contain the 48 active links of a cell of side two, each once.

TABLE IV. The sixteen block plaquettes modulo two: the base point x (digits $x_0 x_1 x_2 x_3$ modulo two) and the plane.

x	plane	x	plane
0000	$\{0, 1\}$	0000	$\{2, 3\}$
0001	$\{0, 1\}$	0011	$\{2, 3\}$
0010	$\{0, 1\}$	0100	$\{2, 3\}$
0011	$\{0, 1\}$	0111	$\{2, 3\}$
1100	$\{0, 1\}$	1000	$\{2, 3\}$
1101	$\{0, 1\}$	1011	$\{2, 3\}$
1110	$\{0, 1\}$	1100	$\{2, 3\}$
1111	$\{0, 1\}$	1111	$\{2, 3\}$

In closed form, the plaquette at x in $\{0, 1\}$ is a block plaquette when $x_0 \equiv x_1$ and the plaquette at x in $\{2, 3\}$ when

$x_2 \equiv x_3$; the blocks do not depend on the line field beyond the fact that every plaquette contains one fixed link. The bounds of Lemma 16 are evaluated in ball arithmetic at 120 bits: the series of Z_0 and of its first three derivatives in $s = r^2$ is summed to forty terms, and the tail is bounded for $s \leq 5$ by twice the first omitted term $(5/4)^{40} 40^3 / (40! 41!)$, because consecutive terms decrease by a factor below one half. The constants of (3) and α are then evaluated at a given β_W from exact input.

For $SU(3)$ the bounds of Lemma 38 are evaluated on the 841 boxes of side $\frac{3}{16}$ at $r = \frac{3}{2}$ and the 784 boxes of side $\frac{19}{100}$ at $r = \frac{38}{25}$ with the moments of Appendix E, and the constants of (15) and α_3 are evaluated from these bounds and the cover of Proposition 34 or Proposition 35 in ball arithmetic at 120 bits. The cover of Proposition 35 is computed as in Appendix E with $r = \frac{38}{25}$ and the threshold 1.03.

Appendix D: The blocks of Lemma 20

Diagonal blocks. Let $w_a = e^{i\phi}$. On the diagonal part, $f_p = -\cos \phi (y_1 + y_2 - y_3 - y_4)^2$ and $c_p = 2(y_1 y_4 - y_1 y_2 + y_2 y_3 - y_3 y_4)$, so

$$\mathcal{F} = \sum_i \alpha_i y_i^2 + \cos \phi (y_1 + y_2 - y_3 - y_4)^2 - \frac{9}{5} (y_1 y_4 - y_1 y_2 + y_2 y_3 - y_3 y_4)$$

with $y_k = 0$ and weights $\alpha_{k+1} = \frac{21}{10}$, $\alpha_{k+2} = \frac{5}{4}$. It is affine in $\cos \phi$, so positivity at $\cos \phi = \pm 1$ suffices. For $k = 1$ and $\cos \phi = 1$ the matrix in (y_2, y_3, y_4) is

$$\begin{pmatrix} \frac{31}{10} & -\frac{19}{10} & -1 \\ -\frac{19}{10} & \frac{9}{4} & \frac{19}{10} \\ -1 & \frac{19}{10} & \frac{31}{10} \end{pmatrix},$$

and Table V lists the leading principal minors for every position; all are positive.

TABLE V. Leading principal minors of the diagonal blocks of Lemma 20.

Fixed link	$\cos \phi = 1$	$\cos \phi = -1$
e_1, e_4	$\frac{31}{10}, \frac{673}{200}, \frac{8421}{2000}$	$\frac{11}{10}, \frac{53}{200}, \frac{21}{2000}$
e_2	$\frac{31}{10}, \frac{861}{100}, \frac{8421}{2000}$	$\frac{11}{10}, \frac{21}{100}, \frac{21}{2000}$
e_3	$\frac{9}{4}, \frac{673}{200}, \frac{8421}{2000}$	$\frac{1}{4}, \frac{53}{200}, \frac{21}{2000}$

Pair blocks. Let $w_a = e^{i(\chi+\vartheta)}$ and $w_b = e^{i(\chi-\vartheta)}$. The terms of f_p carry $w_a + \bar{w}_b = 2\cos \chi e^{i\vartheta}$ or its conjugate, and c_p carries $w_a \bar{w}_b = e^{2i\vartheta}$, so the Hermitian matrix $M(m, \vartheta)$ of $\mathcal{F}$, $m = \cos \chi$, is affine in m with $M(-1, \vartheta) = M(1, \vartheta + \pi)$, and $m = 1$ suffices. For the fixed link e_2 , in the variables (z_1, z_3, z_4) and with $c = \cos \vartheta$,

$$M = \begin{pmatrix} \frac{21}{5} + 2c & -2e^{i\vartheta} & -\frac{9}{5} - 2e^{i\vartheta} \\ -2e^{-i\vartheta} & \frac{21}{5} + 2c & \frac{9}{5} + 2e^{-i\vartheta} \\ -\frac{9}{5} - 2e^{-i\vartheta} & \frac{9}{5} + 2e^{i\vartheta} & \frac{5}{2} + 2c \end{pmatrix}.$$

With $u = \cos^2(\vartheta/2) \in [0, 1]$ its leading principal minors are

$$\frac{20u+11}{5}, \quad \frac{(20u+1)(20u+21)}{25}, \quad \frac{400u(40u^2-23u+4)+21}{250},$$

and $40u^2 - 23u + 4$ has negative discriminant $529 - 640$. The other positions give the same determinant and the leading minors $\frac{20u+11}{5}$ or $\frac{8u+1}{2}$ and $\frac{800u^2-180u+53}{50}$, whose discriminant $32400 - 169600$ is negative. Every block is positive definite.

The inequality is not tight. In floating point the least eigenvalue of the whole form over $W \in U(N)$, $N \leq 4$, is 0.0149, which is also the least eigenvalue of the exact blocks; this slack is what the choice of the weights $\frac{21}{10}, \frac{5}{4}, \frac{9}{10}$ trades against the constant $\frac{29}{2}$.

Appendix E: The certified computation

Truncation. For a corner b of a box the moments of Lemma 32 are derivatives of Z at $(b, \bar{b})$. The truncated series is $\mathbb{E}_H[A_t(UD)B_t(\bar{U}D')]$ with A_t the part of $\sum_\lambda c_\lambda s_\lambda(UD)$ over partitions with $\lambda_1 - \lambda_3 \leq 34$ and $\lambda_3 \leq 11$, and B_t likewise. Every omitted partition has $|\lambda| > 34$, because $3 \cdot 12 > 34$, and the omitted part is at most $T(\|D\|) = \sum_{n>34} (\frac{3}{2}\|D\|)^n / n!$ in modulus. On the polydisc $|D_{ij} - b_{ij}| \leq \rho$, $\|D\| \leq \|b\| + 3\rho$, and Cauchy's estimate bounds the α -th derivative of the omitted part by $\alpha! \rho^{-|\alpha|} T(\|b\| + 3\rho)$. The derivatives of the full exponential are at most $2^{-|\alpha|} e^{v/2}$ with $v = \sum_k |b_k|$, since $|\text{Tr}(Ub)| \leq \sum_k |b_k|$ for diagonal b . With $\varepsilon_\alpha = 2^{|\alpha|} \alpha! \rho^{-|\alpha|} T(\|b\| + 3\rho)$ the truncated moment differs from the true one by at most $\varepsilon_\alpha e^{v/2} + (e^{v/2} + \varepsilon_\alpha) \varepsilon_\beta$, and ρ is chosen to minimise $\rho^{-6} T$. The largest error bound of an unnormalised moment at any corner is 1.7×10^{-11} .

Weight zero. Conjugation by the diagonal torus T preserves Haar measure and $\text{ReTr}(Ub)$ for diagonal b , and multiplies u_{ij} by $t_i t_j^{-1}$; a moment of a monomial on which T acts by a nontrivial character is therefore 0, and only 8278 of the 134596 monomials of degree at most six have the trivial character. The functions $\text{Tr}(X)$, $e_2(X)$, $\det X$ of $X = DD'^\top$ and $\det D$, $\det D'$ are invariant under the same action on the variables, so after these are formed in the full ring of Taylor polynomials of degree six in the 18 variables, the series is computed in the subring of weight-zero monomials.

Rounding. The truncated series is a Taylor polynomial of total degree six in the 36 real directions of $(D - b, D' - \bar{b})$, computed in double precision. Every operation is repeated on absolute values and carries an operation count n , so that the computed coefficient differs from the exact one by at most $\gamma_n = nu/(1 - nu)$, $u = 2^{-53}$, times the absolute-value coefficient [43, Sec. 3.1]. With C in $2^{-30}\mathbb{Z}$, the polynomials $2^{31}q_h$, $X_a(2^{31}q_h)$, $12 \cdot 2^{31}(\varphi_h + \Delta q_h)$ and $12 \cdot 2^{31}\langle \nabla \varphi_{F_k}, \nabla q_h \rangle$ for the coordinate vectors h have coefficients in $\mathbb{Z}[i]$, and every partial sum in their computation is bounded, through the same operations on absolute values, by less than 2^{50} ; they are therefore computed exactly. Normalisation by Z , the products with the corner coordinates, the quadratic forms and their error matrices are formed with the same bookkeeping, and the largest eigenvalues are bounded in ball arithmetic [48].

Separation. The functional c of a discarded box is the dual solution of a linear program for the gauge of the centre of the box with respect to a 72×72 grid of T . Only the final inequality is used: $\max_T \ell$ is bounded by its maximum over a 512×512 grid plus $\frac{1}{2}H\delta^2/2$, with $\delta = 2\pi/512$ and $H = |c_1| + |c_2| + 2|c_3|$ a bound for its Hessian on the torus, and the inequality is checked with relative margin 10^{-12} .

Appendix F: The transfer matrix

This appendix recalls the Osterwalder–Schrader reconstruction for the lattice [62–64] in the form used for the gaps, following our earlier paper [14].

Let ω be a state on the gauge-invariant local functions of $\mathbb{Z}^4$ that is invariant under translations and the hypercubic group and reflection positive for reflections in planes through sites and in planes bisecting links; by Proposition 11(iii)–(iv) and its analogues in the other theorems, the states of this paper are of this kind. The reconstruction gives a Hilbert space with a vector Ω , a map $F \mapsto \Psi_F$ with dense range, and a self-adjoint $\mathbb{T}$ with $0 \leq \mathbb{T} \leq 1$, $\mathbb{T}\Omega = \Omega$ and $\langle \Psi_F, \mathbb{T}'\Psi_H \rangle = \omega(\vartheta(\bar{F})\tau_t H)$, where ϑ is the time reflection and τ_t the time translation.²¹

Proposition 116. *For $m > 0$ the following are equivalent:*

- (a) *the spectrum of $\mathbb{T}$ on $\Omega^\perp$ lies in $[0, e^{-m}]$;*
- (b) *for every F in a class of gauge-invariant local functions supported at times in a bounded interval and dense in $L^2(\omega)$ on each such support, there is $C_F < \infty$ with $0 \leq \omega(\vartheta(\bar{F})\tau_t F) - |\omega(F)|^2 \leq C_F e^{-mt}$ for all $t \geq 0$.*

Under (a) one may take $C_F = \|F\|_\infty^2$ for every bounded F .

Proof. If (a) holds, then

$$\omega(\vartheta(\bar{F})\tau_t F) - |\omega(F)|^2 = \langle \Psi_F', \mathbb{T}'\Psi_F' \rangle$$

with $\Psi_F' = \Psi_F - \langle \Omega, \Psi_F \rangle \Omega \in \Omega^\perp$; this number lies in $[0, e^{-mt}\|\Psi_F'\|^2]$, and $\|\Psi_F'\|^2 \leq \|F\|_\infty^2$. Conversely, let σ be the spectral measure of $\mathbb{T}$ at Ψ_F' and let $\delta > 0$. Then

$$\sigma((e^{-m+\delta}, 1])e^{(-m+\delta)t} \leq C_F e^{-mt}$$

for all t , so $\sigma((e^{-m+\delta}, 1]) = 0$. The spectral projection of $(e^{-m}, 1]$ restricted to $\Omega^\perp$ therefore annihilates the vectors Ψ_F' of the class, which are dense by the footnote above, and it vanishes. $\square$

²¹ Ψ_F is the class of F in the completion of the positive-time local functions for the form $\langle F, H \rangle = \omega(\vartheta(\bar{F})H)$, which is positive semidefinite by reflection positivity. Since $|\omega(\vartheta(\bar{F})F)| \leq \omega(|F|^2)$ by the Cauchy–Schwarz inequality and the invariance of ω under ϑ , functions that are close in $L^2(\omega)$ give close vectors.

Appendix G: Spectral measures and Laplace transforms

This appendix collects the facts on spectral measures that the continuum step uses, with proofs.

Lemma 117. *Let T be a self-adjoint operator on a Hilbert space with $0 \leq T \leq 1$ and let ψ be a unit vector. There is a unique probability measure ν on $[0, 1]$ with $\langle \psi, h(T)\psi \rangle = \int h d\nu$ for every bounded Borel function h on $[0, 1]$. If $T\psi' = \psi'$ implies $\langle \psi', \psi \rangle = 0$, then $\nu(\{1\}) = 0$. If μ is the image of ν under $\lambda \mapsto -c^{-1} \log \lambda \in [0, \infty]$ for some $c > 0$, then $\langle \psi, T^{\sigma/c} \psi \rangle = \int_{[0, \infty]} e^{-\sigma E} d\mu(E)$ for $\sigma > 0$, with the integrand 0 at $E = \infty$.*

Proof. The spectral theorem [65, Sec. VII.2] gives the measure ν , carried by the spectrum of T , which lies in $[0, 1]$; uniqueness holds because the moments $\int \lambda^n d\nu = \langle \psi, T^n \psi \rangle$ determine a measure on a compact interval, polynomials being dense in $C[0, 1]$. The spectral projection $P = \mathbf{1}_{\{1\}}(T)$ projects onto the eigenvectors of T with eigenvalue 1, so $\nu(\{1\}) = \|P\psi\|^2 = \langle P\psi, \psi \rangle = 0$ under the hypothesis. The last identity is the change of variables $\lambda^{\sigma/c} = e^{-\sigma E}$ for $\lambda = e^{-cE}$, with $\lambda = 0$ corresponding to $E = \infty$. $\square$

Lemma 118. *Every sequence of probability measures on a compact metric space K has a subsequence that converges weakly, that is, against every continuous function, to a probability measure on K .*

Proof. $C(K)$ is separable; let (h_j) be dense. The numbers $\int h_j d\mu_k$ are bounded by $\|h_j\|_\infty$, so a diagonal subsequence makes all of them converge. By density the limit $\Lambda(h) = \lim \int h d\mu_k$ exists for every $h \in C(K)$; Λ is linear, positive and $\Lambda(1) = 1$, so by the Riesz representation theorem it is integration against a probability measure. $\square$

Lemma 119. *Let μ be a probability measure on $[0, \infty]$ and $\rho(\sigma) = \int e^{-\sigma E} d\mu(E)$ for $\sigma > 0$.*

- (a) *ρ is nonincreasing and log-convex: $\rho(\sigma)^2 \leq \rho(\sigma - s)\rho(\sigma + s)$ for $0 < s < \sigma$.*
- (b) *$\rho(\sigma) > 0$ for one $\sigma > 0$ if and only if $\mu([0, \infty)) > 0$, and then $\rho(\sigma) > 0$ for every $\sigma > 0$.*
- (c) *$\rho(\sigma) \rightarrow \mu([0, \infty))$ as $\sigma \downarrow 0$.*
- (d) *If μ is carried by $[m, \infty]$, then $\rho(\sigma) \leq e^{-m\sigma}$.*
- (e) *For $M > 0$, $\mu([0, M]) \geq \rho(\sigma) - e^{-\sigma M}$.*
- (f) *If $\mu_k \rightarrow \mu$ weakly on $[0, \infty]$, then $\int e^{-\sigma E} d\mu_k \rightarrow \rho(\sigma)$ for every $\sigma > 0$.*

Proof. (a) The integrand decreases in σ . By the Cauchy–Schwarz inequality,

$$\left(\int e^{-\sigma E} d\mu \right)^2 \leq \int e^{-(\sigma-s)E} d\mu \int e^{-(\sigma+s)E} d\mu.$$

(b) The integrand is positive exactly on $[0, \infty)$. (c) The integrand increases to $\mathbf{1}_{[0, \infty)}$ as $\sigma \downarrow 0$, and monotone convergence

applies. (d) The integrand is at most $e^{-m\sigma}$ on $[m, \infty]$. (e) Split the integral at M : it is at most $\mu([0, M]) + e^{-\sigma M}$. (f) For $\sigma > 0$ the function $E \mapsto e^{-\sigma E}$, extended by 0 at ∞ , is continuous on $[0, \infty]$. $\square$

By (b) and (e), a lower bound $\rho(\sigma) \geq c > 0$ at one σ places mass at least $c/2$ on the bounded interval $[0, \sigma^{-1} \log(2/c)]$; for the measures μ_k of Theorem 48 a bound uniform in k therefore makes them tight on $[0, \infty)$ with a fixed fraction of their mass.

Appendix H: The character expansion

Let G be a compact connected Lie group with normalised Haar measure dU . For an irreducible unitary representation ρ_λ of dimension d_λ write $\chi_\lambda = \text{Tr} \rho_\lambda$. The characters form an orthonormal basis of the square-integrable class functions (Peter–Weyl and Schur orthogonality [22, Chs. 2–4]).

Lemma 120. *Let $w \in L^1(G)$ be a class function with $w(U^{-1}) = w(U)$, and put*

$$\hat{c}_\lambda = d_\lambda^{-1} \int w \chi_\lambda dU.$$

Then $\int w(U) \rho_\lambda(U) dU = \hat{c}_\lambda I$, and $\int w(U) \chi_\lambda(U^{-1}V) dU = \hat{c}_\lambda \chi_\lambda(V)$.

Proof. $M = \int w(U) \rho_\lambda(U) dU$ commutes with every $\rho_\lambda(V)$, since $w(VUV^{-1}) = w(U)$ and Haar measure is invariant; by Schur's lemma $M = cI$, and taking traces gives $c = \hat{c}_\lambda$. Then $\int w(U) \chi_\lambda(U^{-1}V) dU = \text{Tr}(\int w(U) \rho_\lambda(U^{-1}) dU \rho_\lambda(V))$, and the substitution $U \mapsto U^{-1}$ with $w(U^{-1}) = w(U)$ turns the integral into M . $\square$

Lemma 121. *Let $G = \text{SU}(N)$ and $\beta > 0$, and let*

$$w_\beta(U) = e^{\frac{\beta}{N} \text{Re Tr } U} / \int e^{\frac{\beta}{N} \text{Re Tr } V} dV.$$

Then $\hat{c}_\lambda(\beta) = d_\lambda^{-1} \int w_\beta \chi_\lambda dU$ is real, $0 \leq \hat{c}_\lambda(\beta) \leq 1$, and $\hat{c}_\lambda(\beta) < 1$ for $\lambda \neq 0$. Moreover

$$\hat{c}_\lambda(\beta) = 1 - \frac{NC_2(\lambda)}{\beta} + O(\beta^{-2}) \quad (\beta \rightarrow \infty),$$

where $-\sum_a \rho_\lambda(T_a)^2 = 2C_2(\lambda)I$ for a basis (T_a) of $\mathfrak{su}(N)$ orthonormal for $|X|^2 = -\text{Tr } X^2$, so that $C_2 = (N^2 - 1)/(2N)$ on the fundamental representation.

Proof. Since $w_\beta(\bar{U}) = w_\beta(U)$ and $\chi_\lambda(\bar{U}) = \overline{\chi_\lambda(U)}$, and complex conjugation preserves Haar measure, $\int w_\beta \chi_\lambda = \int w_\beta \bar{\chi}_\lambda$, which is the complex conjugate of the left side; so $\hat{c}_\lambda$ is real. Next, w_β is a multiple of $\exp(\frac{\beta}{2N}(\chi_F + \bar{\chi}_F)) = \sum_m \frac{1}{m!} (\frac{\beta}{2N})^m (\chi_F + \bar{\chi}_F)^m$, and each power is the character of a representation, hence a sum of irreducible characters with non-negative integer coefficients. By orthonormality $\int w_\beta \bar{\chi}_\lambda \geq 0$, so $\hat{c}_\lambda \geq 0$. As $|\chi_\lambda(U)| \leq d_\lambda$, with $\text{Re } \chi_\lambda(U) = d_\lambda$ only where $\rho_\lambda(U) = I$, a proper closed subgroup when $\lambda \neq 0$ and so a Haar null set, and $w_\beta > 0$ everywhere, $\hat{c}_\lambda = d_\lambda^{-1} \int w_\beta \text{Re } \chi_\lambda < 1$ for $\lambda \neq 0$.

For the expansion, $\text{Re Tr } U \leq N$ with equality only at $U = I$, so outside a neighbourhood $\{U = e^X : |X| < \delta\}$ both integrals $\int e^{\frac{\beta}{N} \text{Re Tr } U} \chi_\lambda$ and $\int e^{\frac{\beta}{N} \text{Re Tr } U}$ are smaller than their values near I by a factor $e^{-c\beta}$. Near I write $U = e^X$; the Haar density in these coordinates is $j(X) = 1 + O(|X|^2)$, even in X . For antihermitian X , $\text{Tr } X^3$ is imaginary, so

$$\frac{1}{N} \text{Re Tr } e^X = 1 - \frac{1}{2N} |X|^2 + r(X), \quad 0 \leq r(X) \leq C|X|^4,$$

and $d_\lambda^{-1} \chi_\lambda(e^X) = 1 + d_\lambda^{-1} \text{Tr } \rho_\lambda(X) + \frac{1}{2d_\lambda} \text{Tr } \rho_\lambda(X)^2 + c_3(X) + O(|X|^4)$, with c_3 cubic and $\text{Tr } \rho_\lambda(X) = 0$, since $\mathfrak{su}(N)$ is perfect. Let $\langle \cdot \rangle$ denote the average against the Gaussian weight $e^{-\beta|X|^2/(2N)}$, under which the coordinates x_a of $X = \sum x_a T_a$ are independent with variance N/β . With $h = j e^{\beta r}$,

$$\begin{aligned} \hat{c}_\lambda(\beta) &= \frac{\langle d_\lambda^{-1} \chi_\lambda(e^X) h \rangle}{\langle h \rangle} + O(e^{-c\beta}) \\ &= 1 + \frac{\langle \frac{1}{2d_\lambda} \text{Tr } \rho_\lambda(X)^2 h \rangle}{\langle h \rangle} + \frac{\langle (c_3(X) + O(|X|^4)) h \rangle}{\langle h \rangle}, \end{aligned}$$

the averages being taken over $|X| < \delta$. Under $\langle \cdot \rangle$, $|X|^{2m}$ has mean of order β^{-m} and $\beta|r(X)| \leq C\beta|X|^4$ has mean of order β^{-1} , so $h = 1 + O(\beta^{-1})$ in every moment. The weight h is even in X , since $\text{Tr } X^k$ is imaginary for odd k and j is even, so the odd term c_3 contributes nothing; the quartic remainder and the correlation of the quadratic term with $h - 1$ contribute $O(\beta^{-2})$. Finally $\langle \text{Tr } \rho_\lambda(X)^2 \rangle = \frac{N}{\beta} \sum_a \text{Tr } \rho_\lambda(T_a)^2 = -\frac{2NC_2(\lambda)d_\lambda}{\beta}$, which gives the expansion. For $\text{SU}(2)$ and the representation of spin j the integrals are explicit, $\hat{c}_j(\beta) = I_{2j+1}(\beta)/I_1(\beta)$ with the modified Bessel functions I_n , in agreement with $C_2 = j(j+1)$. $\square$

Appendix I: The single-link law of $\text{SU}(2)$

This appendix proves the sharp quarter, $W_1(v_b, v_{b'}) \leq \frac{1}{4}|b - b'|$ for the one-link law of $\text{SU}(2)$, which Theorems 1 and 3 use, and the exact value at $b = 0$, which Theorem 2 uses. The argument is that of our earlier paper [14]; it is given in full so that the present paper is self-contained.

1. The von Mises–Fisher law

Identify $\text{SU}(2)$ with the unit quaternions, that is with the unit sphere $S^3 \subset \mathbb{R}^4$, so that $\text{Re Tr } U = 2x_0(U)$ and $\text{Re Tr}(UV) = 2\langle u, \bar{v} \rangle$ for the Euclidean inner product of $\mathbb{R}^4$ and quaternion conjugation $v \mapsto \bar{v}$. Fix a link e and the field off e . Every plaquette through e can be rotated cyclically so that its trace reads $\text{Re Tr}(Q_e A_p)$ with $A_p \in \text{SU}(2)$, the staple of e in p . The action therefore depends on Q_e only through

$$N\beta \sum_{p \ni e} \text{Re Tr}(Q_e A_p) = \langle x, b \rangle, \quad b = 4\beta \sum_{p \ni e} \bar{A}_p \in \mathbb{R}^4,$$

with $x = Q_e \in S^3$. Since there are six staples, each of norm one,

$$\kappa := |b| \leq y := 24|\beta|. \quad (11)$$

The conditional law of Q_e is the von Mises–Fisher law²²

$$v_b(dx) = Z(\kappa)^{-1} e^{\langle x, b \rangle} \sigma(dx) \quad (I2)$$

on the unit sphere, with σ the normalised surface measure. We use coordinates in which $b = \kappa e_1$. The density of $t = x_1$ is proportional to $(1 - t^2)^{1/2} e^{\kappa t}$ on $[-1, 1]$, so $Z(\kappa) = 2I_1(\kappa)/\kappa$ with the modified Bessel function I_1 , normalised so that $Z(0) = 1$.

Lemma 122. *For every $\kappa \geq 0$ the law v_b on the unit sphere S^3 satisfies the Poincaré inequality $\text{Var}_{v_b}(g) \leq \frac{1}{3} \int |\nabla g|^2 dv_b$, and the constant $\frac{1}{3}$ is attained at $\kappa = 0$.*

Proof. In the polar angle $\theta \in (0, \pi)$ from e_1 the generator of the reversible diffusion of v_b is $L = \partial_\theta^2 + b(\theta)\partial_\theta + \sin^{-2}\theta \Delta_{S^2}$ with $b(\theta) = 2\cot\theta - \kappa \sin\theta$. On spherical harmonics of degree ℓ on S^2 it acts as a Sturm–Liouville operator L_ℓ with potential $\ell(\ell+1)\sin^{-2}\theta$. A ground state comparison of Barta’s type [9] with the trial function $\psi = \sin\theta e^{a\kappa \cos\theta}$ gives $-L_1\psi \geq 3\psi$ for $a = -\frac{1}{5}$, and the same trial function with $a = -\frac{2}{5}$ gives the bound 3 for the derivatives of zonal eigenfunctions; Appendix J carries out both computations. Hence every nonzero eigenvalue of $-L$ is at least 3. At $\kappa = 0$ the first eigenvalue of the Laplacian of the unit S^3 is 3. $\square$

2. Moments

Let $A(\kappa) = \mathbb{E}_{v_b} x_1 = Z'(\kappa)/Z(\kappa) = I_2(\kappa)/I_1(\kappa)$.

Lemma 123. *For every $\kappa > 0$, in coordinates with $b = \kappa e_1$,*

(i) $\mathbb{E}x_2^2 = \mathbb{E}x_3^2 = \mathbb{E}x_4^2 = A/\kappa$, $\mathbb{E}(1 - x_1^2) = 3A/\kappa$, the covariance matrix of x is $\text{diag}(A', A/\kappa, A/\kappa, A/\kappa)$, and $A' = 1 - A^2 - 3A/\kappa$;

$$(ii) \quad \frac{\kappa}{2 + \sqrt{\kappa^2 + 4}} \leq A(\kappa) \leq \frac{\kappa}{1 + \sqrt{\kappa^2 + 9}};$$

$$(iii) \quad A' \leq A/\kappa \leq (1 + \sqrt{\kappa^2 + 9})^{-1} \leq \frac{1}{4};$$

$$(iv) \quad A(\kappa) \geq \kappa / \left(\frac{3}{2} + \sqrt{\kappa^2 + 25/4} \right).$$

Proof. (i) Integrating by parts against $\frac{d}{dt}(1 - t^2)^{3/2} = -3t(1 - t^2)^{1/2}$ gives $\kappa \mathbb{E}(1 - x_1^2) = 3\mathbb{E}x_1$, and the three coordinates x_2, x_3, x_4 are exchangeable under v_b , with $\sum_i x_i^2 = 1$. The off-diagonal covariances vanish by the reflections $x_i \mapsto -x_i$, $i \geq 2$, and $A' = \text{Var}x_1$ because $A = (\log Z)'$. (ii) By (i), A solves the Riccati equation $A' = 1 - A^2 - 3A/\kappa$ with $A(\kappa) =$

$\kappa/4 + O(\kappa^3)$. Appendix J shows that the left side of (ii) is a subsolution and the right side a supersolution of this equation, with the same behaviour at 0, and concludes by comparison. (iii) The middle inequality is (ii). For the first, put $s = \sqrt{\kappa^2 + 4}$. By (ii), $4A/\kappa \geq 4/(2+s)$ and $1 - A^2 \leq 1 - \kappa^2/(2+s)^2$, and

$$1 - \frac{\kappa^2}{(2+s)^2} \leq \frac{4}{2+s} \iff (2+s)^2 - \kappa^2 \leq 4(2+s) \\ \iff s^2 - \kappa^2 \leq 4,$$

which holds with equality. So $A' = 1 - A^2 - 3A/\kappa \leq A/\kappa$. (iv) is the comparison of (ii) with a third barrier, which improves the lower bound of (ii); it is carried out in the same appendix. $\square$

3. The Lipschitz rate of the family

Let δ be the chordal metric $\delta(x, x') = |x - x'|$ on S^3 and W_1 the Kantorovich distance for δ . For $c \in \mathbb{R}^4$ and a unit vector $h \in \mathbb{R}^4$ put

$$L(c, h) = \sup_g \text{Cov}_{v_c}(g, \langle h, x \rangle), \quad (I3)$$

the supremum over 1-Lipschitz functions g for δ . Since $\frac{d}{dt} v_{c+th}(g) = \text{Cov}_{v_{c+th}}(g, \langle h, x \rangle)$, the Kantorovich–Rubinstein duality gives, for all b, b' ,

$$W_1(v_b, v_{b'}) \leq |b - b'| \sup_{c \in [b, b']} L\left(c, \frac{b' - b}{|b' - b|}\right). \quad (I4)$$

By rotation invariance $L(c, h)$ depends only on $\kappa = |c|$ and on the angle φ between h and $c^\perp$, and we write $L(\kappa, \varphi)$.

Lemma 124. *For all c and h , $L(c, h) \leq 1/(2\sqrt{3})$. Moreover:*

(i) at $c = 0$, $L(0, h) = \frac{1}{4}$ for every h ;

(ii) if $h \perp c$, then $L(c, h) = \mathbb{E}_{v_c} x_2^2 = A(\kappa)/\kappa \leq \frac{1}{4}$;

(iii) if $\kappa \geq \sqrt{88}/3 \approx 3.127$, then $L(c, h) \leq \frac{1}{4}$ for every h .

Proof. A 1-Lipschitz function for δ is 1-Lipschitz for the geodesic distance, which is larger, so $|\nabla g| \leq 1$ almost everywhere, and Lemma 122 gives $\text{Var} g \leq \frac{1}{3}$. By Lemma 123(i), (iii) the largest eigenvalue of the covariance of x is $A/\kappa \leq \frac{1}{4}$, so $\text{Var} \langle h, x \rangle \leq A/\kappa$. The Cauchy–Schwarz inequality gives

$$L(c, h) \leq \left(\frac{1}{3} A/\kappa \right)^{1/2} \leq \left(3(1 + \sqrt{\kappa^2 + 9}) \right)^{-1/2} \leq \frac{1}{2\sqrt{3}}. \quad (I5)$$

(i) and (ii) are reflection arguments. Let R be the reflection in the hyperplane $h^\perp$, and suppose v_c is R -invariant, which is the case when $c = 0$ or $h \perp c$. Then $\mathbb{E} \langle h, x \rangle = 0$ and

$$\text{Cov}(g, \langle h, x \rangle) = \mathbb{E}[(g(x) - g(Rx))\langle h, x \rangle; \langle h, x \rangle > 0] \\ \leq \mathbb{E}[2\langle h, x \rangle^2; \langle h, x \rangle > 0] = \mathbb{E} \langle h, x \rangle^2,$$

since $\delta(x, Rx) = 2|\langle h, x \rangle|$. Equality holds for $g = \langle h, x \rangle$, which is 1-Lipschitz. At $c = 0$, $\mathbb{E} \langle h, x \rangle^2 = \frac{1}{4}$; for $h \perp c$, $\mathbb{E} \langle h, x \rangle^2 = A/\kappa$. (iii) By (I5), $L \leq \frac{1}{4}$ as soon as $1 + \sqrt{\kappa^2 + 9} \geq \frac{16}{3}$, that is $\kappa^2 \geq \frac{88}{9}$. $\square$

²² On the sphere $S^{p-1} \subset \mathbb{R}^p$ the von Mises–Fisher law with mean direction μ and concentration κ has density proportional to $e^{\kappa \langle x, \mu \rangle}$ with respect to the surface measure, and its normalising constant is $(2\pi)^{p/2} I_{p/2-1}(\kappa)/\kappa^{p/2-1}$; it is the maximum entropy law with prescribed mean resultant, and it plays on the sphere the role of the Gaussian in $\mathbb{R}^p$. For $p = 4$ the Bessel index is 1. The mean resultant length $\mathbb{E} \langle x, \mu \rangle = I_{p/2}(\kappa)/I_{p/2-1}(\kappa)$ is the function A of Lemma 123. See Mardia and Jupp [58], Ch. 9, for these facts and for the tangent-normal decomposition used in Appendix J.

The reflection settles the directions $h \perp c$ and the point $c = 0$; for an oblique h the reflection no longer preserves v_c , and the Cauchy–Schwarz bound (15) is too weak for $\kappa < \sqrt{88}/3$. The next theorem closes this case.

Theorem 125 (The sharp quarter). *For every $c \in \mathbb{R}^4$ and every unit vector $h \in \mathbb{R}^4$,*

$$L(c, h) \leq \frac{1}{4},$$

with equality at $c = 0$. Equivalently, $W_1(v_b, v_{b'}) \leq \frac{1}{4}|b - b'|$ for all $b, b' \in \mathbb{R}^4$, and the constant $\frac{1}{4}$ is the best possible.

The equivalence is (14) in one direction and the derivative formula above it in the other; the sharpness is Lemma 124(i). The metric matters.²³ The proof has four steps. A Lipschitz function on S^3 extends to the unit ball $B \subset \mathbb{R}^4$, and Green’s formula in B bounds the covariance by a negative Sobolev norm of the density (Lemma 126). This norm splits into a zonal and a transverse part, both explicit series in modified Bessel functions (Lemma 127). The transverse part is at most $\frac{1}{4}$ because a certain power series has nonnegative coefficients (Proposition 129), and the zonal part is at most $\frac{1}{4}$ by a Riccati comparison (Proposition 130).

4. Harmonic extension

Let σ be the normalised surface measure of S^3 and H_k the space of spherical harmonics of degree k , the restrictions to S^3 of harmonic homogeneous polynomials of degree k on $\mathbb{R}^4$. The spaces H_k are orthogonal in $L^2(\sigma)$ and span it; we write u_k for the component of u in H_k , and for u with $\int u d\sigma = 0$ we put

$$\mathcal{Q}(u) = \sum_{k \geq 1} \frac{1}{k} \|u_k\|_{L^2(\sigma)}^2. \quad (\text{I6})$$

This is $\langle u, \Lambda^{-1}u \rangle$ for the Dirichlet-to-Neumann operator Λ of B , which acts on H_k as multiplication by k .

Lemma 126. *Let u be a real-analytic function on S^3 with $\int u d\sigma = 0$ and let g be 1-Lipschitz for the chordal metric. Then*

$$\int_{S^3} g u d\sigma \leq \frac{1}{2} \mathcal{Q}(u)^{1/2}.$$

Proof. The chordal metric is the Euclidean metric of $\mathbb{R}^4$ restricted to S^3 , so McShane’s extension $G(y) = \inf_{x \in S^3} (g(x) + |x - y|)$ is 1-Lipschitz on $\mathbb{R}^4$ and equals g on S^3 [59]; by

²³ In the geodesic metric of S^3 the statement is false already at $c = 0$: the function $\arcsin x_2$ is geodesically 1-Lipschitz and its covariance with x_2 under the uniform law is $\frac{8}{9\pi} \approx 0.283 > \frac{1}{4}$. A proof therefore has to use that the chordal distance between two points is shorter than their geodesic distance, that is, it has to see the sphere from the ball it bounds. The proof below does exactly this: it extends a chordally Lipschitz function to the ball and integrates by parts there.

Rademacher’s theorem $|\nabla G| \leq 1$ almost everywhere. Let $\psi = \sum_{k \geq 1} k^{-1} r^k u_k(\omega)$ in polar coordinates $y = r\omega$. Each term is a harmonic polynomial, the series converges with all derivatives on the closed ball because u is real-analytic,²⁴ and $\partial_r \psi = u$ on S^3 . Green’s formula for the Lipschitz function G and the smooth harmonic function ψ reads $\int_{S^3} G \partial_r \psi dS = \int_B \nabla G \cdot \nabla \psi dy$. With $|S^3| = 2\pi^2$ and $|B| = \pi^2/2$ the Cauchy–Schwarz inequality gives

$$\int g u d\sigma = \frac{1}{2\pi^2} \int_B \nabla G \cdot \nabla \psi dy \leq \frac{|B|^{1/2}}{2\pi^2} \left(\int_B |\nabla \psi|^2 dy \right)^{1/2}.$$

Green’s formula once more gives $\int_B |\nabla \psi|^2 = \int_{S^3} \psi \partial_r \psi dS = 2\pi^2 \sum_k k^{-1} \|u_k\|^2 = 2\pi^2 \mathcal{Q}(u)$, and the right side becomes $\frac{1}{2} \mathcal{Q}(u)^{1/2}$. $\square$

The lemma is a form of Thomson’s principle. By Kantorovich–Rubinstein duality, $\sup_g \int g u d\sigma$ is the least cost of a flow in $\mathbb{R}^4$ that carries the positive part of $u d\sigma$ to its negative part, the cost being the total length travelled by the mass.²⁵ The harmonic field $\nabla \psi$ is one such flow, through the interior of the ball, and the lemma computes its cost. At $c = 0$ the flow is optimal: for $u = \langle h, x \rangle$ one has $\psi = \langle h, y \rangle$, the field $\nabla \psi = h$ is constant, the Cauchy–Schwarz inequality is an equality, and $\mathcal{Q}(u) = \|u\|^2 = \frac{1}{4}$.

5. Spectral decomposition

Fix $c = \kappa e_1$ with $\kappa > 0$ and write $h = \cos \varphi e_1 + \sin \varphi e_2$, which is no loss of generality by rotation invariance. Let $\rho = dv_c/d\sigma = e^{\kappa x_1}/Z(\kappa)$, $Z(\kappa) = \int e^{\kappa x_1} d\sigma = 2I_1(\kappa)/\kappa$, and

$$u = (\langle h, x \rangle - \mathbb{E}_{v_c} \langle h, x \rangle) \rho,$$

so that $\text{Cov}_{v_c}(g, \langle h, x \rangle) = \int g u d\sigma$ and $\int u d\sigma = 0$. Let U_n be the Chebyshev polynomials of the second kind and put

$$d_n(\kappa) = \frac{(n+1)I_{n+1}(\kappa)}{I_1(\kappa)}, \quad n \geq 0. \quad (\text{I7})$$

Lemma 127. *With this notation*

$$\mathcal{Q}(u) = \cos^2 \varphi \mathcal{Q}_z(\kappa) + \sin^2 \varphi \mathcal{Q}_x(\kappa),$$

where

$$\mathcal{Q}_z = \sum_{n \geq 1} \frac{d'_n(\kappa)^2}{n}, \quad \mathcal{Q}_x = \frac{1}{3\kappa^2} \sum_{n \geq 1} (n+2) d_n(\kappa)^2.$$

²⁴ For real-analytic u the norms $\|u_k\|$ decay geometrically in k , and the supremum of a function in H_k is at most $(\dim H_k)^{1/2} = k+1$ times its $L^2(\sigma)$ norm, so the series for ψ and for each of its derivatives converge uniformly on $\bar{B}$.

²⁵ This is the Beckmann formulation of the transport problem with cost $|x - y|$: the optimal cost equals the least $\int |V|$ over vector measures V on $\mathbb{R}^4$ with divergence $u d\sigma$. Restricting the flow to the closed ball does not change the infimum, since projecting onto $\bar{B}$ shortens paths. The gradient $\nabla \psi$ on B is one admissible flow. It minimises the energy $\int |V|^2$ rather than the cost $\int |V|$, and the Cauchy–Schwarz inequality converts one into the other at the price of the factor $|B|^{1/2}$.

Proof. The functions $U_n(x_1)$, $n \geq 0$, are the zonal harmonics: $U_n(x_1) \in H_n$, and they are orthonormal in $L^2(\sigma)$ (Appendix K). The expansion $e^{\kappa t} = \frac{2}{\kappa} \sum_n (n+1) I_{n+1}(\kappa) U_n(t)$ of the same appendix gives $\rho = \sum_n d_n(\kappa) U_n(x_1)$, with $d_0 = 1$.

Write $u = \cos \varphi u_z + \sin \varphi u_x$ with $u_z = (x_1 - A)\rho$ and $u_x = x_2 \rho$; here $\mathbb{E}_{v_c} x_1 = A$ and $\mathbb{E}_{v_c} x_2 = 0$. The reflection $x_2 \mapsto -x_2$ maps each H_k to itself, fixes u_z and reverses u_x , so the components of u_z and u_x in each H_k are orthogonal and the cross term in $\mathcal{Q}(u)$ vanishes.

Since $A = Z'/Z$, we have $u_z = \partial_\kappa \rho = \sum_n d'_n(\kappa) U_n(x_1)$, and the component of u_z in H_n is $d'_n U_n(x_1)$, of norm $|d'_n|$. This gives $\mathcal{Q}_z$.

For u_x , let $y(t) = \cos t e_1 + \sin t e_2$. The law $v_{\kappa y(t)}$ has density $\rho_t = e^{\kappa \langle y(t), x \rangle} / Z(\kappa) = \sum_n d_n(\kappa) U_n(\langle y(t), x \rangle)$, and $\partial_t \rho_t|_{t=0} = \kappa x_2 \rho$. The function $U_n(\langle y(t), x \rangle)$ lies in H_n for every t , hence so does its derivative $x_2 U'_n(x_1)$ at $t = 0$, and the component of u_x in H_n is $\kappa^{-1} d_n x_2 U'_n(x_1)$. Its norm follows from the eigenvalue $n(n+2)$ of the Laplacian of S^3 on H_n . For $f = U_n(x_1)$,

$$\begin{aligned} n(n+2) &= n(n+2) \|f\|^2 = \int |\nabla_{S^3} f|^2 d\sigma \\ &= \int U'_n(x_1)^2 (1 - x_1^2) d\sigma, \end{aligned}$$

since the tangential gradient of x_1 has squared length $1 - x_1^2$. Given x_1 , the point (x_2, x_3, x_4) is uniform on the sphere of radius $(1 - x_1^2)^{1/2}$, so $\mathbb{E}[x_2^2 | x_1] = \frac{1}{3}(1 - x_1^2)$ and $\|x_2 U'_n(x_1)\|^2 = \frac{1}{3} n(n+2)$. Summing $n^{-1} \kappa^{-2} d_n^2 \cdot \frac{1}{3} n(n+2)$ over $n \geq 1$ gives $\mathcal{Q}_x$. $\square$

By Lemmas 126 and 127,

$$L(\kappa, \varphi) \leq \frac{1}{2} \max(\mathcal{Q}_z(\kappa), \mathcal{Q}_x(\kappa))^{1/2}, \quad (I8)$$

and Theorem 125 follows once both parts are at most $\frac{1}{4}$. Both are $\frac{1}{4}$ at $\kappa = 0$, by continuity from the values $d_1 \sim \kappa/2$, $d'_1 \rightarrow \frac{1}{2}$ and $d_n = O(\kappa^n)$.

6. Three Bessel sums

The series of Lemma 127 reduce to closed forms. Write I_j for $I_j(\kappa)$.

Lemma 128. *For every $\kappa > 0$,*

$$\begin{aligned} \sum_{j \geq 1} j^2 I_j^2 &= \frac{\kappa}{4} I_1(2\kappa), & \sum_{j \geq 1} j^3 I_j^2 &= \frac{\kappa^2}{4} (I_0^2 + I_1^2), \\ I_1(2\kappa) &= 2 \sum_{j \geq 1} I_{j-1} I_j. \end{aligned}$$

The first and third identities follow from Parseval's formula for $e^{\kappa \cos \theta} = \sum_{j \in \mathbb{Z}} I_j e^{ij\theta}$, and the second from the recurrences of I_j by telescoping; Appendix K gives the proofs.

7. The transverse part

Proposition 129. $\mathcal{Q}_x(\kappa) \leq \frac{1}{4}$ for every $\kappa \geq 0$.

Proof. By (I7), $\mathcal{Q}_x = (3\kappa^2 I_1^2)^{-1} \sum_{j \geq 2} j^2 (j+1) I_j^2$, and the sum equals $\sum_{j \geq 1} (j^3 + j^2) I_j^2 - 2I_1^2$. By Lemma 128, $\mathcal{Q}_x \leq \frac{1}{4}$ is equivalent to

$$\mathcal{D}(\kappa) := (8 + 2\kappa^2) I_1^2 - \kappa^2 I_0^2 - \kappa I_1(2\kappa) \geq 0. \quad (I9)$$

Each term is an entire function of κ that is even after the factor κ^2 is removed, so $\mathcal{D}(\kappa) = \sum_{n \geq 0} \mathcal{D}_n \kappa^{2n}$. Lemma 132 of Appendix K computes the coefficients: $\mathcal{D}_0 = \mathcal{D}_1 = 0$, and for $n = m+1 \geq 2$

$$\mathcal{D}_n = \frac{N_m}{4^m m! (m+1)! (m+1)(m+2)},$$

with

$$\begin{aligned} N_m &= \binom{2m}{m} (4(2m+1) + (m-1)(m+1)(m+2)) \\ &\quad - 4^m (m+1)(m+2). \end{aligned}$$

The numerator N_m vanishes for $m = 0, 1, 2$ and is positive for $m \geq 3$ (Lemma 132). So every coefficient of $\mathcal{D}$ is nonnegative and $\mathcal{D} \geq 0$. $\square$

The first nonzero coefficient is $\mathcal{D}_4 = \frac{1}{2304}$, so $\frac{1}{4} - \mathcal{Q}_x = \mathcal{D}/(12\kappa^2 I_1^2)$ is of order κ^4 at 0: the transverse part leaves the value $\frac{1}{4}$ very slowly, which is why no estimate that loses a constant can prove the proposition.

8. The zonal part

Let $A = I_2/I_1$ as before, so that $d_1 = 2A$ and, by Lemma 123(i), $A' = 1 - A^2 - 3A/\kappa$.

Proposition 130. *For $0 < \kappa \leq \sqrt{88}/3$,*

$$\mathcal{Q}_z(\kappa) \leq \frac{1}{4} + 4A'^2 + \frac{A^2}{4} - \frac{A}{2\kappa} - \frac{2A^2}{\kappa^2} < \frac{1}{4}.$$

Proof. The higher modes. With $r_v = I_{v+1}/I_v$ the recurrence $I'_v = I_{v+1} + \frac{v}{\kappa} I_v$ gives

$$\frac{d'_n}{d_n} = \frac{I'_{n+1}}{I_{n+1}} - \frac{I'_1}{I_1} = \frac{n}{\kappa} + r_{n+1} - r_1.$$

By Lemma 133 of Appendix K, $r_{n+1} \leq r_1$, and $r_1 = A \leq \kappa/4$ by Lemma 123(iii). Hence $-n/\kappa \leq n/\kappa - \kappa/4 \leq d'_n/d_n \leq n/\kappa$ for $n \geq 2$ and $\kappa^2 \leq 16$, so $d_n^2 \leq n^2 d_n'^2 / \kappa^2$. With $d'_1 = 2A'$,

$$\mathcal{Q}_z \leq 4A'^2 + \frac{1}{\kappa^2} \sum_{n \geq 2} n d_n'^2 = 4A'^2 + \frac{S_3 - S_2 - 4I_2^2}{\kappa^2 I_1^2},$$

where $S_2 = \sum_{j \geq 1} j^2 I_j^2$ and $S_3 = \sum_{j \geq 1} j^3 I_j^2$; the second form uses $n d_n'^2 I_1^2 = (j-1) j^2 I_j^2$ with $j = n+1$.

The closed form. The recurrences $I_0 - I_2 = \frac{2}{\kappa} I_1$ and $I_1 - I_3 = \frac{4}{\kappa} I_2$ give $I_0/I_1 = 2/\kappa + A$ and $I_2 I_3 / I_1^2 = A - 4A^2/\kappa$. By

Lemma 128, $S_3/(\kappa^2 I_1^2) = \frac{1}{4}(2/\kappa + A)^2 + \frac{1}{4}$, and, keeping three terms of the third identity,

$$\frac{S_2}{\kappa^2 I_1^2} \geq \frac{I_0 I_1 + I_1 I_2 + I_2 I_3}{2\kappa I_1^2} = \frac{1}{2\kappa} \left(\frac{2}{\kappa} + 3A - \frac{4A^2}{\kappa} \right).$$

Collecting terms, $\mathcal{Q}_z \leq \frac{1}{4} + 4A^2 + T(A)$ with $T(a) = \frac{a^2}{4} - \frac{a}{2\kappa} - \frac{2a^2}{\kappa^2}$, which is the first inequality.

The barrier. Let $s = \sqrt{\kappa^2 + 25/4}$ and $A_b = \kappa/(\frac{3}{2} + s)$. By Lemma 123(iv), $A_b \leq A \leq \kappa/4$. The function $a \mapsto 1 - a^2 - 3a/\kappa$ decreases on $a \geq 0$, and $A' \geq 0$ because $A' = \text{Var} x_1$, so $0 \leq A' \leq \alpha := 1 - A_b^2 - 3A_b/\kappa$. For $\kappa \leq 12$ the quadratic T decreases on $[0, \kappa/4]$, since

$$T(a) - T(a') = (a' - a) \frac{2\kappa - (a + a')(\kappa^2 - 8)}{4\kappa^2}$$

and $(a + a')(\kappa^2 - 8) \leq \frac{\kappa}{2}(\kappa^2 - 8) \leq 2\kappa$ there. So $\mathcal{Q}_z \leq \frac{1}{4} + 4\alpha^2 + T(A_b)$. With $\kappa^2 = s^2 - \frac{25}{4}$ and $A_b/\kappa = 2/(2s + 3)$ both terms are rational in s : $\alpha = 16/(2s + 3)^2$, and

$$4\alpha^2 + T(A_b) = \frac{(2s - 5)p(s)}{4(2s + 3)^4}, \quad (I10)$$

$$p(s) = 8s^3 + 28s^2 - 98s - 695.$$

For $0 < \kappa \leq \sqrt{88}/3$ we have $\frac{5}{2} < s \leq \sqrt{577}/6 < 4.004$. On $s \geq \frac{5}{2}$ the cubic p increases, since $p'(s) = 24s^2 + 56s - 98 > 0$, and $p(4.004) < -120$. So the right side of (I10) is negative. $\square$

9. Proof of Theorem 125

At $c = 0$ the theorem is Lemma 124(i), and for $|c| \geq \sqrt{88}/3$ it is Lemma 124(iii). For $0 < |c| < \sqrt{88}/3$, Propositions 129 and 130 give $\mathcal{Q}_x, \mathcal{Q}_z \leq \frac{1}{4}$, and (I8) gives $L \leq \frac{1}{4}$. $\square$

Appendix J: Computations for the single-link law

1. Barta's inequality

Let $\rho(\theta) = \sin^2 \theta e^{\kappa \cos \theta}$ on $(0, \pi)$, let V be a continuous potential, and let

$$T = -\rho^{-1} \partial_\theta \rho \partial_\theta + V.$$

If $\psi > 0$ is smooth on $(0, \pi)$ with $\rho \psi^2 = O(\sin^4 \theta)$ at the end points and $T\psi \geq \lambda_0 \psi$, then

$$\int (|h'|^2 + V h^2) \rho d\theta \geq \lambda_0 \int h^2 \rho d\theta$$

for every smooth h vanishing to first order at 0 and π . Indeed, writing $h = \psi u$ and integrating by parts, the boundary terms $\rho \psi \psi' u^2$ vanish and

$$\int (|h'|^2 + V h^2) \rho = \int \rho \psi^2 |u'|^2 + \int \frac{T\psi}{\psi} h^2 \rho \geq \lambda_0 \int h^2 \rho.$$

This is the inequality of Barta [9] in the form of a ground state transformation.

2. The computations of Lemma 122

The generator of the reversible diffusion of v_b on S^3 is $L = \Delta_{S^3} + \kappa \nabla x_1 \cdot \nabla$. In the polar angle θ from e_1 the round metric is $d\theta^2 + \sin^2 \theta g_{S^2}$, so $\Delta_{S^3} = \partial_\theta^2 + 2 \cot \theta \partial_\theta + \sin^{-2} \theta \Delta_{S^2}$, and $\nabla x_1 \cdot \nabla = -\sin \theta \partial_\theta$. On $h(\theta) Y_\ell$ with a spherical harmonic Y_ℓ of degree ℓ on S^2 the operator acts as $L_\ell = \partial_\theta^2 + b \partial_\theta - \ell(\ell + 1) \sin^{-2} \theta$, $b = 2 \cot \theta - \kappa \sin \theta = \rho'/\rho$, so $-L_\ell = T$ with $V = \ell(\ell + 1) \sin^{-2} \theta$.

Degree $\ell \geq 1$. Take $\psi = \sin \theta e^{a \kappa \cos \theta}$. Then $\psi'/\psi = \cot \theta - a \kappa \sin \theta$ and $\psi''/\psi = (\psi'/\psi)^2 + (\psi'/\psi)'$, and a short computation gives

$$\frac{-L_1 \psi}{\psi} = 3 + (1 + 5a) \kappa \cos \theta - a(1 + a) \kappa^2 \sin^2 \theta.$$

For $a = -\frac{1}{5}$ this is $3 + \frac{4}{25} \kappa^2 \sin^2 \theta \geq 3$. Since $V = \ell(\ell + 1) \sin^{-2} \theta \geq 2 \sin^{-2} \theta$ for $\ell \geq 1$, Barta's inequality gives $-L_\ell \geq 3$ on every sector with $\ell \geq 1$.

Degree $\ell = 0$. Let $g(\theta)$ be a zonal eigenfunction, $-L_0 g = \lambda g$, g not constant, so that $g'(0) = g'(\pi) = 0$. Differentiating $g'' + b g' = -\lambda g$ gives, for $h = g'$,

$$-(h'' + b h') - b' h = \lambda h, \quad -b' = \frac{2}{\sin^2 \theta} + \kappa \cos \theta,$$

so h is an eigenfunction of T with $V = 2 \sin^{-2} \theta + \kappa \cos \theta$, with eigenvalue λ , and $h \not\equiv 0$ vanishes at both ends. With the same ψ ,

$$\frac{T\psi}{\psi} = 3 + (2 + 5a) \kappa \cos \theta - a(1 + a) \kappa^2 \sin^2 \theta,$$

which for $a = -\frac{2}{5}$ is $3 + \frac{6}{25} \kappa^2 \sin^2 \theta \geq 3$. Barta's inequality applied to h gives $\lambda \geq 3$. Together with the case $\ell \geq 1$ every nonzero eigenvalue of $-L$ is at least 3, which is Lemma 122; the spectrum is discrete because S^3 is compact and the drift is smooth.

3. Moments

The density of $t = x_1$ under v_b is proportional to $(1 - t^2)^{1/2} e^{\kappa t}$. For every smooth ϕ on $[-1, 1]$,

$$\frac{d}{dt} \left[(1 - t^2)^{3/2} e^{\kappa t} \right] = (1 - t^2)^{1/2} e^{\kappa t} (\kappa(1 - t^2) - 3t),$$

and integration by parts gives $\mathbb{E}[\phi'(t)(1 - t^2)] = \mathbb{E}[\phi(t)(3t - \kappa(1 - t^2))]$. With $\phi = 1$ this is $\kappa \mathbb{E}(1 - x_1^2) = 3 \mathbb{E} x_1$, the identity of Lemma 123(i). With $\phi = t^k$ it gives the recursion

$$m_{k+2} = m_k - \frac{k+3}{\kappa} m_{k+1} + \frac{k}{\kappa} m_{k-1}, \quad m_k = \mathbb{E} x_1^k,$$

from $m_0 = 1$ and $m_1 = A$; all polynomial moments of v_b are rational in κ and $A(\kappa)$.²⁶

²⁶ For instance $m_2 = 1 - 3A/\kappa$, $m_3 = A - 3/\kappa + 12A/\kappa^2$ and $m_4 = m_2 - 5m_3/\kappa + 2A/\kappa$. The joint moments with x_2, x_3, x_4 follow from exchangeability: $\mathbb{E}[\phi(x_1)x_2^2] = \frac{1}{3} \mathbb{E}[\phi(x_1)(1 - x_1^2)]$.

4. The Riccati barriers

By Lemma 123(i), A solves $A' = \mathcal{F}(\kappa, A) := 1 - A^2 - 3A/\kappa$ on $(0, \infty)$, and $A(\kappa) = \frac{\kappa}{4} - \frac{\kappa^3}{96} + O(\kappa^5)$ from the series of I_1 and I_2 . Let $s = \sqrt{\kappa^2 + 4}$, $u = \sqrt{\kappa^2 + 9}$, $\phi_- = \kappa/(2 + s)$ and $\phi_+ = \kappa/(1 + u)$. Direct differentiation gives

$$\begin{aligned}\phi'_- - \mathcal{F}(\kappa, \phi_-) &= -\frac{\kappa^2}{s(s+2)^2} \leq 0, \\ \phi'_+ - \mathcal{F}(\kappa, \phi_+) &= \frac{(u-3)^2}{u(1+u)^2} \geq 0,\end{aligned}$$

and $\phi_{\pm} = \frac{\kappa}{4} + O(\kappa^3)$.²⁷ The comparison is singular at 0 because of the term $3A/\kappa$, and we carry it out with an integrating factor. Let $D = A - \phi_-$. Then

$$\begin{aligned}D' &= \mathcal{F}(\kappa, A) - \mathcal{F}(\kappa, \phi_-) - (\phi'_- - \mathcal{F}(\kappa, \phi_-)) \\ &= -\left(A + \phi_- + \frac{3}{\kappa}\right)D + R,\end{aligned}$$

with $R = \kappa^2/(s(s+2)^2) \geq 0$. Put $\mu(\kappa) = \kappa^3 \exp \int_0^\kappa (A + \phi_-)$. Then $(\mu D)' = \mu R \geq 0$, and $\mu D = O(\kappa^6) \rightarrow 0$ as $\kappa \rightarrow 0$, so $\mu D \geq 0$ and $A \geq \phi_-$. The same argument with $D = \phi_+ - A$ and $R = (u-3)^2/(u(1+u)^2)$ gives $A \leq \phi_+$. This is Lemma 123(ii).

For (iv) let $s = \sqrt{\kappa^2 + 25/4}$ and $\phi_b = \kappa/(\frac{3}{2} + s)$. Since $s^2 - \frac{9}{4} = \kappa^2 + 4$, the same differentiation gives

$$\phi'_b - \mathcal{F}(\kappa, \phi_b) = -\frac{5(s - \frac{5}{2})}{2s(\frac{3}{2} + s)^2} \leq 0,$$

so ϕ_b is a subsolution, and $\phi_b = \frac{\kappa}{4} - \frac{\kappa^3}{80} + O(\kappa^5)$.²⁸ Hence $A - \phi_b = \frac{\kappa^3}{480} + O(\kappa^5)$, and the integrating-factor argument above, with $D = A - \phi_b$, $R = \frac{5(s-5/2)}{2s(\frac{3}{2}+s)^2} \geq 0$ and $\mu = \kappa^3 \exp \int_0^\kappa (A + \phi_b)$, gives $\mu D = O(\kappa^6)$ at 0, $(\mu D)' = \mu R \geq 0$, and so $A \geq \phi_b$. Since $\frac{3}{2} + \sqrt{\kappa^2 + 25/4} < 2 + \sqrt{\kappa^2 + 4}$ for $\kappa > 0$, this improves the lower bound of (ii).

Appendix K: Bessel functions

This appendix proves the auxiliary statements used in Appendix I for Theorem 125. Throughout, $I_j = I_j(\kappa)$ with $\kappa > 0$,

²⁷ For the first identity write $\phi_- = (s-2)/\kappa$, which follows from $s^2 - 4 = \kappa^2$; then $\phi'_- = (\kappa^2/s - s + 2)/\kappa^2$ and $1 - \phi_-^2 - 3\phi_-/\kappa = (\kappa^2 - (s-2)^2 - 3(s-2))/\kappa^2$, and the difference simplifies with $s^2 = \kappa^2 + 4$. The second is the same computation with $\phi_+ = \kappa(u-1)/(u^2-1) = \kappa(u-1)/(\kappa^2+8)$ in place of ϕ_- and $u^2 = \kappa^2 + 9$ in place of $s^2 = \kappa^2 + 4$. After clearing the denominator $\kappa^2(1+u)^2$ the numerator becomes $\kappa^2(u-3)^2/u$, which is the stated residual.

²⁸ With $\phi_b = \kappa(s - \frac{3}{2})/(\kappa^2 + 4)$ one has $\phi'_b = (\frac{3}{2}s + \frac{25}{4})/(s(\frac{3}{2} + s)^2)$, using $ds/d\kappa = \kappa/s$; and $\mathcal{F}(\kappa, \phi_b) = 1 - \phi_b^2 - 3\phi_b/\kappa$. Clearing the denominator $s(\frac{3}{2} + s)^2$ and replacing κ^2 by $s^2 - \frac{25}{4}$ leaves $-\frac{5}{2}s + \frac{25}{4}$ in the numerator, which is the stated residual. The expansion follows from $s = \frac{5}{2} + \frac{\kappa^2}{5} + O(\kappa^4)$.

and $I_{-j} = I_j$ for integers j . We use the series

$$I_\nu(\kappa) = \sum_{i \geq 0} \frac{(\kappa/2)^{2i+\nu}}{i!(i+\nu)!},$$

the integral $I_j(\kappa) = \pi^{-1} \int_0^\pi e^{\kappa \cos \theta} \cos j\theta d\theta$, the Fourier series $e^{\kappa \cos \theta} = \sum_{j \in \mathbb{Z}} I_j e^{ij\theta}$ and the recurrences

$$I_{j-1} - I_{j+1} = \frac{2j}{\kappa} I_j, \quad I'_j = I_{j+1} + \frac{j}{\kappa} I_j, \quad (\text{K1})$$

all of which are in Watson's treatise [76], §§3.7, 3.71, 2.22. All series below converge absolutely, because $0 < I_j \leq (\kappa/2)^j e^\kappa / j!$.

1. Zonal harmonics

Under σ the coordinate x_1 has the density $\frac{2}{\pi}(1-t^2)^{1/2}$ on $[-1, 1]$, the area of the two-sphere $\{x_1 = t\}$ being proportional to $1-t^2$ and its normal speed to $(1-t^2)^{-1/2}$. The Chebyshev polynomials $U_n(\cos \theta) = \sin((n+1)\theta)/\sin \theta$ satisfy

$$\int U_n U_m d\sigma = \frac{2}{\pi} \int_0^\pi \sin((n+1)\theta) \sin((m+1)\theta) d\theta = \delta_{nm},$$

and $U_n = C_n^{(1)}$ is the Gegenbauer polynomial of index $\frac{4-2}{2} = 1$, so $U_n(\langle y, x \rangle)$ is the zonal harmonic of degree n on S^3 with pole y : $|x|^n U_n(\langle y, x \rangle / |x|)$ is a harmonic polynomial on $\mathbb{R}^4$ [72], Ch. IV, §2.

Lemma 131. $e^{\kappa t} = \frac{2}{\kappa} \sum_{n \geq 0} (n+1) I_{n+1}(\kappa) U_n(t)$ in $L^2(\frac{2}{\pi}(1-t^2)^{1/2} dt)$, and uniformly on $[-1, 1]$.

Proof. With $t = \cos \theta$ and $2 \sin \theta \sin((n+1)\theta) = \cos n\theta - \cos(n+2)\theta$, the n -th coefficient is

$$\begin{aligned}\frac{2}{\pi} \int_0^\pi e^{\kappa \cos \theta} \sin((n+1)\theta) \sin \theta d\theta &= I_n - I_{n+2} \\ &= \frac{2(n+1)}{\kappa} I_{n+1},\end{aligned}$$

by (K1). Since $|U_n| \leq n+1$ and I_{n+1} decays factorially, the series converges uniformly. $\square$

2. Proof of Lemma 128

Let $F(\theta) = e^{\kappa \cos \theta} = \sum_j I_j e^{ij\theta}$. Then

$$F'(\theta) = -\kappa \sin \theta F(\theta) = \sum_j ij I_j e^{ij\theta},$$

and Parseval's formula gives

$$\begin{aligned}\sum_{j \in \mathbb{Z}} j^2 I_j^2 &= \frac{\kappa^2}{2\pi} \int_0^{2\pi} \sin^2 \theta e^{2\kappa \cos \theta} d\theta \\ &= \frac{\kappa^2}{2} (I_0(2\kappa) - I_2(2\kappa)) = \frac{\kappa}{2} I_1(2\kappa),\end{aligned}$$

the last step by (K1) at the argument 2κ . Halving gives the first identity. The coefficient of $e^{i\theta}$ in $F^2 = e^{2\kappa \cos \theta}$ is $I_1(2\kappa) = \sum_{j \in \mathbb{Z}} I_j I_{1-j}$, and the terms with $j \leq 0$ repeat those with $j \geq 1$; this is the third identity.

For the second, $jI_j = \frac{\kappa}{2}(I_{j-1} - I_{j+1})$ by (K1), so

$$\sum_{j \geq 1} j^3 I_j^2 = \frac{\kappa^2}{4} \sum_{j \geq 1} j(I_{j-1}^2 + I_{j+1}^2 - 2I_{j-1}I_{j+1}).$$

Shifting indices, $\sum_{j \geq 1} j(I_{j-1}^2 + I_{j+1}^2) = I_0^2 + \sum_{m \geq 1} 2mI_m^2$, and $\sum_{m \geq 1} 2mI_m^2 = \kappa \sum_{m \geq 1} (I_{m-1}I_m - I_mI_{m+1}) = \kappa I_0I_1$ by telescoping. For the mixed term write $j = \frac{1}{2}((j-1) + (j+1))$ and apply (K1) to $(j-1)I_{j-1}$ and to $(j+1)I_{j+1}$, using $I_{-1} = I_1$ at $j = 1$:

$$jI_{j-1}I_{j+1} = \frac{\kappa}{4} [(I_{j-1}I_j - I_jI_{j+1}) + (I_{j-2}I_{j+1} - I_{j-1}I_{j+2})].$$

Both brackets telescope, to I_0I_1 and to $I_{-1}I_2 = I_1I_2$, so $\sum_{j \geq 1} jI_{j-1}I_{j+1} = \frac{\kappa}{4}(I_0I_1 + I_1I_2)$. Altogether

$$\sum_{j \geq 1} j^3 I_j^2 = \frac{\kappa^2}{4} (I_0^2 + \frac{\kappa}{2} I_1(I_0 - I_2)) = \frac{\kappa^2}{4} (I_0^2 + I_1^2). \quad \square$$

3. The coefficients of $\mathcal{D}$

Lemma 132. Write $\mathcal{D}(\kappa) = \sum_{n \geq 0} \mathcal{D}_n \kappa^{2n}$ for the function of (19). Then $\mathcal{D}_0 = \mathcal{D}_1 = 0$, and for $m \geq 1$

$$\mathcal{D}_{m+1} = \frac{N_m}{4^m m! (m+1)! (m+1)(m+2)},$$

with $N_m = \binom{2m}{m} (4(2m+1) + (m-1)(m+1)(m+2)) - 4^m(m+1)(m+2)$. Moreover $N_m = 0$ for $m \leq 2$ and $N_m > 0$ for $m \geq 3$.

Proof. The coefficients. Multiplying the series of I_0 and I_1 and collecting powers,

$$I_0^2 = \sum_{m \geq 0} \alpha_m \kappa^{2m}, \quad I_1^2 = \sum_{m \geq 0} \beta_m \kappa^{2m+2},$$

$$\kappa I_1(2\kappa) = \sum_{m \geq 0} \frac{\kappa^{2m+2}}{m! (m+1)!},$$

where

$$\alpha_m = \frac{1}{4^m} \sum_{i=0}^m \frac{1}{i!^2 (m-i)!^2},$$

$$\beta_m = \frac{1}{4^{m+1}} \sum_{i=0}^m \frac{1}{i! (i+1)! (m-i)! (m-i+1)!}.$$

By Vandermonde's identity, $\sum_i \binom{m}{i}^2 = \binom{2m}{m}$ and $\sum_i \binom{m}{i} \binom{m+2}{i+1} = \binom{2m+2}{m+1}$, so

$$\alpha_m = \frac{\binom{2m}{m}}{4^m m!^2}, \quad \beta_m = \frac{\binom{2m+2}{m+1}}{4^{m+1} m! (m+2)!}.$$

The coefficient of κ^{2m+2} in $\mathcal{D}$ is $\mathcal{D}_{m+1} = 8\beta_m + 2\beta_{m-1} - \alpha_m - 1/(m!(m+1)!)$, with $\beta_{-1} = 0$; for $m = 0$ it is $\frac{8}{4} - 1 - 1 = 0$, and $\mathcal{D}_0 = 0$ because every term of $\mathcal{D}$ is $O(\kappa^2)$. For $m \geq 1$ put $\Delta = 4^m m! (m+1)! (m+1)(m+2)$ and $b = \binom{2m}{m}$. With $\binom{2m+2}{m+1} = \frac{2(2m+1)}{m+1} b$,

$$8\beta_m = \frac{4(2m+1)b}{\Delta}, \quad \alpha_m = \frac{(m+1)^2(m+2)b}{\Delta},$$

$$2\beta_{m-1} = \frac{2m(m+1)(m+2)b}{\Delta},$$

$$\frac{1}{m!(m+1)!} = \frac{4^m(m+1)(m+2)}{\Delta},$$

and $2m - (m+1) = m-1$ gives $\mathcal{D}_{m+1} = N_m/\Delta$.

The sign. Directly, $N_0 = N_1 = N_2 = 0$ ($2-2, 24-24, 192-192$), $N_3 = 80$, $N_4 = 1140$ and $N_5 = 10416$. For $m \geq 6$ we first show $16^m \leq 4m \binom{2m}{m}^2$ for $m \geq 1$ by induction: it is an equality at $m = 1$, and $(m+1) \binom{2m+2}{m+1} = 2(2m+1) \binom{2m}{m}$ gives

$$4(m+1) \binom{2m+2}{m+1}^2 = \frac{16(2m+1)^2}{m+1} \binom{2m}{m}^2$$

$$\geq \frac{16(2m+1)^2}{4m(m+1)} 16^m \geq 16^{m+1}.$$

Since $(m-1)^2 \geq 4m$ for $m \geq 6$, it follows that $\binom{2m}{m}^2 (m-1)^2 \geq 16^m$, that is $\binom{2m}{m} (m-1) \geq 4^m$, and

$$N_m = \left(\binom{2m}{m} (m-1) - 4^m \right) (m+1)(m+2)$$

$$+ 4(2m+1) \binom{2m}{m} > 0.$$

□

4. Monotonicity of Bessel ratios

Lemma 133. For $n \geq 1$ and $\kappa > 0$, $I_{n+2}(\kappa)/I_{n+1}(\kappa) \leq I_2(\kappa)/I_1(\kappa)$.

Proof. We use a lemma of Biernacki and Krzyż [19]: if $f = \sum a_j x^j$ and $g = \sum b_j x^j$ have positive coefficients and a_j/b_j is nonincreasing in j , then f/g is nonincreasing on $x > 0$. Indeed

$$f'g - fg' = \sum_{i,j} (i-j) a_i b_j x^{i+j-1}$$

$$= \frac{1}{2} \sum_{i,j} (i-j) b_i b_j \left(\frac{a_i}{b_i} - \frac{a_j}{b_j} \right) x^{i+j-1} \leq 0.$$

Take $x = \kappa^2$, $f = I_{n+1}(\kappa)/\kappa^{n+1}$ and $g = I_1(\kappa)/\kappa$. Then $a_j/b_j = 2^{-n}(j+1)!/(j+n+1)!$ decreases in j , so $I_{n+1}/(\kappa^n I_1)$ is nonincreasing in κ . Its logarithmic derivative is, by the second recurrence of (K1),

$$\frac{n+1}{\kappa} + \frac{I_{n+2}}{I_{n+1}} - \frac{n}{\kappa} - \frac{1}{\kappa} - \frac{I_2}{I_1}$$

$$= \frac{I_{n+2}}{I_{n+1}} - \frac{I_2}{I_1} \leq 0.$$

□

- [1] A. Adhikari and S. Cao, Correlation decay for finite lattice gauge theories at weak coupling, *Ann. Probab.* **53** (2025), 140–174, doi:10.1214/24-AOP1702.
- [2] G. Aumann, Über Räume mit Mittelbildungen, *Math. Ann.* **119** (1944), 210–215, doi:10.1007/BF01563741.
- [3] D. Bakry and M. Émery, Diffusions hypercontractives, in *Séminaire de Probabilités XIX*, Lecture Notes in Math. 1123, Springer, Berlin, 1985, 177–206, doi:10.1007/BFb0075847.
- [4] D. Bakry, I. Gentil and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, Grundlehren Math. Wiss. 348, Springer, Cham, 2014, doi:10.1007/978-3-319-00227-9.
- [5] T. Balaban, Propagators and renormalization transformations for lattice gauge theories. I, *Comm. Math. Phys.* **95** (1984), 17–40, doi:10.1007/BF01215753.
- [6] T. Balaban, Averaging operations for lattice gauge theories, *Comm. Math. Phys.* **98** (1985), 17–51, doi:10.1007/BF01211042.
- [7] T. Balaban, Large field renormalization. II. Localization, exponentiation, and bounds for the $\mathbf{R}$ operation, *Comm. Math. Phys.* **122** (1989), 355–392, doi:10.1007/BF01238433.
- [8] S. Banach, Über homogene Polynome in (L^2) , *Studia Math.* **7** (1938), 36–44, doi:10.4064/sm-7-1-36-44.
- [9] J. Barta, Sur la vibration fondamentale d’une membrane, *C. R. Acad. Sci. Paris* **204** (1937), 472–473.
- [10] R. Basu and S. Ganguly, $SO(N)$ lattice gauge theory, planar and beyond, *Comm. Pure Appl. Math.* **71** (2018), 2016–2064, doi:10.1002/cpa.21788.
- [11] D. Bhattacharjee, A coherent approach towards quantum gravity, *Phys. Sci. Int. J.* **26**(6) (2022), 59–78, doi:10.9734/psij/2022/v26i6751.
- [12] D. Bhattacharjee, BLLM: Beyond the LLM with Deep Human-Mind-AI, Zenodo, release v3.0.5, 2026, doi:10.5281/zenodo.23469389.
- [13] D. Bhattacharjee, BLLM: Beyond the LLM with Deep Human-Mind-AI, Zenodo, release v3.1.0, 2026, doi:10.5281/zenodo.23484160.
- [14] D. Bhattacharjee and U. Bhattacharya, Mass gap of lattice Yang–Mills theory up to the Dobrushin ceiling, in *Lattice Yang–Mills theory: spacing-uniform obstructions, the mass gap up to the Dobrushin ceiling, and Lean certificate*, Zenodo, release v2.0.2, 2026, doi:10.5281/zenodo.23449466.
- [15] D. Bhattacharjee and U. Bhattacharya, Spacing-uniform lattice inequalities cannot produce the Yang–Mills mass gap, Zenodo, release v1.0.0, 2026, doi:10.5281/zenodo.23444339.
- [16] D. Bhattacharjee, S. Singha Roy and P. Samal, Brane-cluster UV completion of quantum gravity, *Research Square* (2026), doi:10.21203/rs.3.rs-9053205/v1.
- [17] D. Bhattacharjee, P. Samal and S. Bhattacharya, Geometric structure and renormalization group flow in chiral Yang–Mills–Higgs theory, *Preprints* (2026), doi:10.20944/preprints202602.2022.v1.
- [18] D. Bhattacharjee, S. Nath Thakur, S. Bhattacharya and P. Samal, $SU(n)$ holonomy deviations in Calabi–Yau manifolds (CY1–CY4), *Preprints* (2026), doi:10.20944/preprints202602.0023.v1.
- [19] M. Biernacki and J. Krzyż, On the monotonicity of certain functionals in the theory of analytic functions, *Ann. Univ. Mariae Curie-Skłodowska Sect. A* **9** (1955), 135–147.
- [20] M. Biskup, Reflection positivity and phase transitions in lattice spin models, in *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Math. 1970, Springer, Berlin, 2009, 1–86, doi:10.1007/978-3-540-92796-9_1.
- [21] D. Brydges, J. Fröhlich and E. Seiler, On the construction of quantized gauge fields. I. General results, *Ann. Physics* **121** (1979), 227–284, doi:10.1016/0003-4916(79)90098-8.
- [22] D. Bump, *Lie Groups*, 2nd ed., Grad. Texts in Math. 225, Springer, New York, 2013, doi:10.1007/978-1-4614-8024-2.
- [23] S. Cao, R. Nissim and S. Sheffield, Expanded regimes of area law for lattice Yang–Mills theories, preprint, 2025, arXiv:2505.16585.
- [24] S. Chatterjee, Yang–Mills for probabilists, in *Probability and Analysis in Interacting Physical Systems*, Springer Proc. Math. Stat. 283, Springer, Cham, 2019, 1–16, doi:10.1007/978-3-030-15338-0_1.
- [25] S. Chatterjee, Rigorous solution of strongly coupled $SO(N)$ lattice gauge theory in the large N limit, *Comm. Math. Phys.* **366** (2019), 203–268, doi:10.1007/s00220-019-03353-3.
- [26] M. Creutz, Gauge fixing, the transfer matrix, and confinement on a lattice, *Phys. Rev. D* **15** (1977), 1128–1136, doi:10.1103/PhysRevD.15.1128.
- [27] J. M. Combes and L. Thomas, Asymptotic behaviour of eigenfunctions for multiparticle Schrödinger operators, *Comm. Math. Phys.* **34** (1973), 251–270, doi:10.1007/BF01646473.
- [28] J. Dimock, The renormalization group according to Balaban, I. Small fields, *Rev. Math. Phys.* **25** (2013), 1330010, doi:10.1142/S0129055X13300100.
- [29] R. L. Dobrushin, The description of a random field by means of conditional probabilities and conditions of its regularity, *Theory Probab. Appl.* **13** (1968), 197–224, doi:10.1137/1113026.
- [30] R. L. Dobrushin, Prescribing a system of random variables by conditional distributions, *Theory Probab. Appl.* **15** (1970), 458–486, doi:10.1137/1115049.
- [31] R. L. Dobrushin and S. B. Shlosman, Constructive criterion for the uniqueness of Gibbs field, in *Statistical Physics and Dynamical Systems*, Birkhäuser, Boston, 1985, 347–370, doi:10.1007/978-1-4899-6653-7_20.
- [32] B. K. Driver, YM_2 : continuum expectations, lattice convergence, and lassos, *Comm. Math. Phys.* **123** (1989), 575–616, doi:10.1007/BF01218586.
- [33] B. Eckmann, Räume mit Mittelbildungen, *Comment. Math. Helv.* **28** (1954), 329–340, doi:10.1007/BF02566939.
- [34] W. Fulton and J. Harris, *Representation Theory: A First Course*, Grad. Texts in Math. 129, Springer, New York, 1991 (reprinted 2004), doi:10.1007/978-1-4612-0979-9.
- [35] J. Fröhlich, R. Israel, E. H. Lieb and B. Simon, Phase transitions and reflection positivity. I. General theory and long range lattice models, *Comm. Math. Phys.* **62** (1978), 1–34, doi:10.1007/BF01940327.
- [36] J. Fröhlich and T. Spencer, Massless phases and symmetry restoration in abelian gauge theories and spin systems, *Comm. Math. Phys.* **83** (1982), 411–454, doi:10.1007/BF01213610.
- [37] C. Gatteringer and C. B. Lang, *Quantum Chromodynamics on the Lattice*, Lecture Notes in Phys. 788, Springer, Berlin, 2010, doi:10.1007/978-3-642-01850-3.
- [38] R. Ghor, D. Bhattacharjee, P. Nandi, P. Samal, R. Patra, S. Singha Roy and S. Bhattacharya, O’Neill tensor bounds for Riemannian submersions in fibred Calabi–Yau manifolds, *Preprints* (2026), doi:10.20944/preprints202602.0125.v1.
- [39] J. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View*, 2nd ed., Springer, New York, 1987, doi:10.1007/978-1-4612-4728-9.
- [40] L. Gross, C. King and A. Sengupta, Two dimensional Yang–Mills theory via stochastic differential equations, *Ann. Physics* **194** (1989), 65–112, doi:10.1016/0003-4916(89)90032-8.
- [41] A. H. Guth, Existence proof of a nonconfining phase in four-dimensional $U(1)$ lattice gauge theory, *Phys. Rev. D* **21** (1980), 2291–2307, doi:10.1103/PhysRevD.21.2291.

- [42] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Grad. Stud. Math. 34, Amer. Math. Soc., Providence, RI, 2001, doi:10.1090/gsm/034.
- [43] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002, doi:10.1137/1.9780898718027.
- [44] R. Holley and D. Stroock, Logarithmic Sobolev inequalities and stochastic Ising models, *J. Stat. Phys.* **46** (1987), 1159–1194, doi:10.1007/BF01011161.
- [45] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30, doi:10.1080/01621459.1963.10500830.
- [46] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Grad. Texts in Math. 9, Springer, New York, 1972, doi:10.1007/978-1-4612-6398-2.
- [47] A. Jaffe and E. Witten, Quantum Yang–Mills theory, in *The Millennium Prize Problems*, Clay Mathematics Institute and American Mathematical Society, Cambridge, MA and Providence, RI, 2006, 129–152, <https://www.claymath.org/wp-content/uploads/2022/06/yangmills.pdf>.
- [48] F. Johansson, Arb: efficient arbitrary-precision midpoint-radius interval arithmetic, *IEEE Trans. Comput.* **66** (2017), 1281–1292, doi:10.1109/TC.2017.2690633.
- [49] V. G. Kac, *Infinite-Dimensional Lie Algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990, doi:10.1017/CBO9780511626234.
- [50] O. Kallenberg, *Foundations of Modern Probability*, 3rd ed., Springer, Cham, 2021, doi:10.1007/978-3-030-61871-1.
- [51] J. Kogut and L. Susskind, Hamiltonian formulation of Wilson’s lattice gauge theories, *Phys. Rev. D* **11** (1975), 395–408, doi:10.1103/PhysRevD.11.395.
- [52] T. Lévy, Yang–Mills measure on compact surfaces, *Mem. Amer. Math. Soc.* **166** (2003), no. 790, doi:10.1090/memo/0790.
- [53] J. M. Lockwood, Toward the continuum limit of forest-gauge lattice Yang–Mills theory: constrained cubes, exact Gaussian diagnostics, and a gauge-suited mixing test, technical report, 29 October 2026, distributed with the archive.
- [54] M. Lüscher, Construction of a selfadjoint, strictly positive transfer matrix for Euclidean lattice gauge theories, *Comm. Math. Phys.* **54** (1977), 283–292, doi:10.1007/BF01614090.
- [55] M. Lüscher, Properties and uses of the Wilson flow in lattice QCD, *J. High Energy Phys.* **2010**, no. 8, 071, doi:10.1007/JHEP08(2010)071.
- [56] M. Lüscher and P. Weisz, Perturbative analysis of the gradient flow in non-abelian gauge theories, *J. High Energy Phys.* **2011**, no. 2, 051, doi:10.1007/JHEP02(2011)051.
- [57] J. Magnen, V. Rivasseau and R. Sénéor, Construction of YM_4 with an infrared cutoff, *Comm. Math. Phys.* **155** (1993), 325–383, doi:10.1007/BF02097397.
- [58] K. V. Mardia and P. E. Jupp, *Directional Statistics*, Wiley, Chichester, 2000, doi:10.1002/9780470316979.
- [59] E. J. McShane, Extension of range of functions, *Bull. Amer. Math. Soc.* **40** (1934), 837–842, doi:10.1090/S0002-9904-1934-05978-0.
- [60] J. Milnor, Curvatures of left invariant metrics on Lie groups, *Adv. Math.* **21** (1976), 293–329, doi:10.1016/S0001-8708(76)80002-3.
- [61] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge University Press, Cambridge, 1994, doi:10.1017/CBO9780511470783.
- [62] K. Osterwalder and R. Schrader, Axioms for Euclidean Green’s functions, *Comm. Math. Phys.* **31** (1973), 83–112, doi:10.1007/BF01645738.
- [63] K. Osterwalder and R. Schrader, Axioms for Euclidean Green’s functions. II, *Comm. Math. Phys.* **42** (1975), 281–305, doi:10.1007/BF01608978.
- [64] K. Osterwalder and E. Seiler, Gauge field theories on a lattice, *Ann. Physics* **110** (1978), 440–471, doi:10.1016/0003-4916(78)90039-8.
- [65] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, revised and enlarged ed., Academic Press, New York, 1980.
- [66] P. Samal, D. Bhattacharjee and R. Ghora, Renormalization of spacetime dimension, Preprints (2026), doi:10.20944/preprints202602.0531.v1.
- [67] K. Sakai and S. Sasaki, Glueball spectroscopy in lattice QCD using gradient flow, *Phys. Rev. D* **107** (2023), 034510, doi:10.1103/PhysRevD.107.034510.
- [68] E. Seiler, *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics*, Lecture Notes in Phys. 159, Springer, Berlin, 1982, doi:10.1007/3-540-11559-5.
- [69] H. Shen, R. Zhu and X. Zhu, A stochastic analysis approach to lattice Yang–Mills at strong coupling, *Comm. Math. Phys.* **400** (2023), 805–851, doi:10.1007/s00220-022-04609-1.
- [70] R. Slansky, Group theory for unified model building, *Phys. Rep.* **79** (1981), 1–128, doi:10.1016/0370-1573(81)90092-2.
- [71] R. P. Stanley, *Enumerative Combinatorics, Volume 2*, with an appendix by S. Fomin, Cambridge Stud. Adv. Math. 62, Cambridge University Press, Cambridge, 1999, doi:10.1017/CBO9780511609589.
- [72] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, NJ, 1971, doi:10.1515/9781400883899.
- [73] E. Sussman and collaborators, Rotation invariance from lattice symmetry and reflection positivity, manuscript, 2026, in the repository github.com/ethanwsussman/Yang--Mills-LLM-experiment, commit 1b92bd0.
- [74] E. Sussman and collaborators, $SU(2)$ Yang–Mills theory in four dimensions via Balaaban’s renormalisation group: construction, axioms and non-triviality, zeroth draft, 6 October 2026, in the repository github.com/ethanwsussman/Yang--Mills-LLM-experiment, commit 1b92bd0.
- [75] C. Villani, *Optimal Transport: Old and New*, Grundlehren Math. Wiss. 338, Springer, Berlin, 2009, doi:10.1007/978-3-540-71050-9.
- [76] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge, 1944.
- [77] K. G. Wilson, Confinement of quarks, *Phys. Rev. D* **10** (1974), 2445–2459, doi:10.1103/PhysRevD.10.2445.