THE YANG–MILLS MASS GAP ON THE LATTICE: A CONDITIONAL REDUCTION VIA WITTEN LAPLACIAN AND CONSTRUCTIVE RENORMALIZATION

LLUIS ERIKSSON

Abstract. We reduce the weak-coupling lattice Yang–Mills mass gap to four explicitly stated hypotheses: assuming them, $\mathrm{SU}(N_c)$ lattice Yang–Mills theory in $d = 4$ dimensions with Wilson action at sufficiently weak coupling has a positive mass gap $m_{\text{gap}} \geq c(N_c)\, e^{-C(N_c)/g^2} > 0$ in lattice units, uniformly in lattice sizes $L \leq C_0 e^{C/g^2}$. The argument combines Balaban's constructive renormalization group, a Morse–Bott/Witten-Laplacian semiclassical spectral gap estimate at the terminal scale, and a transfer-matrix trace identity. Version 1 of this paper presented the result as self-contained modulo Balaban's RG. Version 2 corrects this assessment: the proof is *conditional* on four explicitly stated hypotheses (Section 1.3). In particular: (i) the Morse–Bott non-degeneracy required by the Helffer–Sjöstrand theory *fails* on the orbifold locus of the flat-connection moduli space — including the minimum $\theta = 0$ of the Born–Oppenheimer potential — where $(d-1)(N_c^2 - 1 - r)$ quartic "toron" zero modes appear, as we exhibit numerically on a real lattice (Proposition 5.3); and (ii) the constants of Balaban's construction must satisfy a quantitative compatibility window $\kappa > C' N_c^{3/2}/\gamma^2$ together with $\gamma^2 \leq 2N_c h_0$, which is empty for typical $O(1)$ decay constants and is not known to follow from Balaban's papers. Version 2 also corrects the sign of the coupling flow in the statement of Balaban's theorem, an inverted extraction regime in the transfer-matrix doubling argument (replaced by a finite-window extraction with explicit error), the definition and Hessian normalization of the Born–Oppenheimer potential (whose v1 form is numerically non-positive), and the Ricci constant of the orbit space ($N_c/2$, not $N_c/4$; direction favorable). All corrections and the surviving ingredients are verified in a companion numerical suite. None of this yields an unconditional result, and the continuum, infinite-volume problem remains expressly out of scope.

Contents

Date: February 2026. Version 2: July 2026.

1. Introduction and main result

1.1. **The problem.** The Yang–Mills mass gap problem [14] asks whether $\mathrm{SU}(N_c)$ gauge theory has a strictly positive mass gap. On the lattice $\Lambda = (\mathbb{Z}/L\mathbb{Z})^4$ with Wilson action, this is the spectral gap of the transfer matrix.

1.2. **Main result.**

Theorem 1.1 (Lattice Yang–Mills mass gap, conditional)**.** *Assume Hypotheses* (H-BAL)*,* (H-CONST)*,* (H-MB) *and* (H-HAM) *of Section* 1.3*. For* $\mathrm{SU}(N_c)$ *on* $(\mathbb{Z}/L\mathbb{Z})^4$ *with Wilson action at* $g^2 \le g_0^2$ *(sufficiently small):*

$$m_{\text{gap}} \;\ge\; c(N_c)\cdot e^{-C(N_c)/g^2} \;>\; 0 \tag{1}$$

uniformly in $L_0 \le L \le C_0 e^{C/g^2}$*, where* $C(N_c) = 24\pi^2 \ln 2/(11N_c)$ *and* $c(N_c) = c_1 N_c^{3/2}\omega_{\min}/\gamma^2 \cdot e^{-O(1)} > 0$*, with* γ *Balaban's fixed terminal coupling,* $\omega_{\min} = \sqrt{N_c S_1}$ *determined by the Born–Oppenheimer potential Hessian (Lemma 5.2; in v1 this read* $\sqrt{N_c S_1/2}$*, corrected here — see Remark 5.4), and* $c_1 > 0$ *depending on* C_0*.*

1.3. **Standing hypotheses (new in v2).** Version 1 presented Theorem 1.1 as self-contained modulo Balaban's renormalization group. The analysis behind version 2 (including the numerical verification suite described in Appendix A) identified the following places where the v1 argument imports unproved statements or fails as written. We state them as explicit hypotheses; Theorem 1.1 is proved *conditionally on all four*.

(H-BAL) Balaban package. The effective actions of Theorem 3.1 exist with the stated properties. Theorem 3.1 is a package paraphrase of the results of [1, 2, 3, 4, 5, 6]; those papers prove ultraviolet stability via renormalization transformations with small-field/large-field decompositions, and the clean form stated here (exact partition-function invariance, single exponentially decaying correction ϵ_k, terminal coupling γ) is an idealization whose precise derivation from [1, 2, 3, 4, 5, 6] is not carried out in this paper.

(H-CONST) Compatibility of constants. The constants of the construction admit the window

$$\gamma_{\min} := \left(\frac{C' N_c^{3/2}}{\kappa}\right)^{1/2} \;<\; \gamma \;\le\; \gamma_{\max} := \sqrt{2N_c h_0}, \tag{W}$$

where κ is Balaban's decay constant (Theorem 3.1(b)), C' is the upper-bound constant of Remark 7.3, and $h_0 \ll 1$ is small enough for the semiclassical expansion of Theorem 6.1 (which requires $h = \gamma^2/(2N_c) \le h_0$). The window (W) is nonempty iff $\kappa > C'\sqrt{N_c}/(2h_0)$ — about $7\,C'$ for $N_c = 2$, $h_0 = 0.1$ — while κ is a *fixed* output of Balaban's expansion, not a parameter that can be taken large. v1 (Remark 7.3) asserted the window is "satisfiable for κ sufficiently large (as guaranteed by Balaban's construction)"; we know of no such guarantee, and we therefore treat (W) as an open quantitative hypothesis. The doubling argument additionally needs $m_k L_k \le \kappa/2$ along the flow (Theorem 7.2), which we fold into (H-CONST).

(H-MB) Semiclassics across the orbifold locus. The Morse–Bott non-degeneracy hypotheses of the Helffer–Sjöstrand/Le Peutrec–Nier–Viterbo theory [11, 12] hold only on the *smooth stratum* of the critical manifold: at points of $\mathcal{M}_{\mathrm{flat}}$ with enhanced stabilizer — the orbifold locus, which contains the global minimum $\theta = 0$ of the Born–Oppenheimer potential — the normal Hessian of the Wilson potential acquires $(d-1)(N_c^2 - 1 - r)$ additional zero modes along which the potential is *quartic* (Proposition 5.3; these are the toron modes of the small-volume analyses [15, 16]). (H-MB) is the assumption that the low-lying cluster structure and the gap $\lambda_1^{\mathrm{eff}} = \sqrt{h}\,\omega_{\min}(1 + o(1))$ of Theorem 6.1 survive this quartic degeneracy. This is plausible — the quartic modes are expected to contribute eigenvalue spacings of order $h^{4/3} \cdot O(1)$ in the relevant normalization, and the small-volume effective Hamiltonians of [15, 16] handle precisely this sector — but it is *not* proved here, and it is not covered by [12], whose hypotheses fail at $\theta = 0$.

(H-HAM) Hamiltonian form at the terminal scale. The transfer-matrix generator of the terminal-scale effective theory is $H = (\gamma/2)(-\Delta_{\mathcal{B}}) + V_{\mathrm{eff}}$ with $V_{\mathrm{eff}} = \beta_{n_{\max}}(V_W + O(e^{-\kappa}))$, as used in Theorem 6.1. The transfer matrix of a Euclidean lattice theory equals e^{-aH} with H of this form only in a time-continuum/anisotropic limit; the identification at the isotropic terminal scale is assumed, not derived from Balaban's effective action.

1.4. **What is new in version 2.** Beyond the honesty retag (title, abstract, hypotheses above), v2 corrects: the sign and normalization of the coupling flow in Theorem 3.1(c) (Remark 3.2); the definition (3) of the Born–Oppenheimer potential, whose v1 form is numerically *non-positive* (Remark 5.4); the Hessian constant in (4) ($N_c S_1$, not $N_c S_1/2$); the extraction regime in the proof of Theorem 7.2, which in v1 was inverted and incompatible with its own error control (Remark 7.4); and the Ricci constant (Remark 4.3). It adds Proposition 5.3 (Morse–Bott failure at the orbifold locus, with a numerical demonstration on a real lattice), the relation to the companion programme papers (Section 9.4), and the verification suite (Appendix A). All v1 numbered statements and equations (1)–(7) retain their numbers; new material is numbered after the v1 items in each section.

2. Setup

2.1. **Lattice gauge theory.** We work on $\Lambda = (\mathbb{Z}/L\mathbb{Z})^d$ with $d = 4$. The connection space is $\mathcal{A} = \prod_{\ell \in \Lambda_1} \mathrm{SU}(N_c)$ with the product bi-invariant metric $\langle X, Y \rangle = -\operatorname{tr}(XY)$. The gauge group $\mathcal{G} = \prod_{x \in \Lambda_0} \mathrm{SU}(N_c)$ acts by $(g \cdot U)_{\ell = (x,\mu)} = g(x) U_\ell g(x + \hat{\mu})^{-1}$.

Definition 2.1. *The gauge orbit space is $\mathcal{B} = \mathcal{A}/\mathcal{G}$ with the quotient metric.*

The Wilson action is $S_W(U) = \sum_{\square}\big(1 - \frac{1}{N_c}\mathrm{Re}\operatorname{tr} U_\square\big)$.

2.2. **Orbifold structure of the orbit space.**

Proposition 2.2. *The orbit space $\mathcal{B} = \mathcal{A}/\mathcal{G}$ is a compact Riemannian orbifold. The singular stratum $\mathcal{B}_{\mathrm{sing}} = \{[U] \in \mathcal{B} : \mathrm{Stab}(U) \supsetneq Z(\mathrm{SU}(N_c))\}$ (reducible connections) has codimension $\ge 2(N_c - 1)$ in $\mathcal{B}$. The smooth stratum $\mathcal{B}^* = \mathcal{B} \setminus \mathcal{B}_{\mathrm{sing}}$ is open, dense, and carries the structure of a smooth Riemannian manifold on which $\pi : \pi^{-1}(\mathcal{B}^*) \to \mathcal{B}^*$ is a Riemannian submersion.*

Proof. $\mathcal{A}$ is a product of compact Lie groups, hence compact, and $\mathcal{G}$ is a compact group acting smoothly by isometries. By the slice theorem, $\mathcal{B} = \mathcal{A}/\mathcal{G}$ is a Riemannian orbifold.

For $\mathrm{SU}(N_c)$ with $N_c \geq 2$: the generic stabilizer is the center $Z(\mathrm{SU}(N_c)) \cong \mathbb{Z}/N_c\mathbb{Z}$. A connection U has larger stabilizer iff there exists $g \notin Z(\mathrm{SU}(N_c))$ commuting with all holonomies of U. This imposes at least $2(N_c - 1)$ independent conditions (the dimension of $\mathrm{SU}(N_c)/T^{N_c-1}$ for the smallest non-central stabilizer). Hence $\mathcal{B}_{\mathrm{sing}}$ has codimension $\geq 2(N_c - 1) \geq 2$ in $\mathcal{B}$. $\square$

3. Balaban's constructive RG

Theorem 3.1 (Balaban [1, 2, 3, 4, 5, 6]; see (H-BAL))**.** *For $\beta = 2N_c/g^2 \geq \beta_0$: effective actions $S_k = \beta_k S_W + \sum_X \epsilon_k(X)$ exist on $\Lambda_k = (\mathbb{Z}/(L/2^k)\mathbb{Z})^4$, $0 \leq k \leq n_{\max}$, with:*

(a) *$Z_0 = Z_k$ for all k and all L divisible by $2^{n_{\max}}$.*
(b) *$|\epsilon_k(X)| \leq e^{-\kappa\,\mathrm{diam}(X)/a_k}$, $\kappa > 0$ universal.*
(c) *$g_k^2 = 2N_c/\beta_k \leq \gamma$ at $k = n_{\max}$, where the coupling flows by asymptotic freedom toward the infrared:*

$$\frac{1}{g_k^2} = \frac{1}{g^2} - 2b_0\,k\ln 2 + O(1), \qquad \textit{i.e.} \qquad \beta_k = \beta - 4N_c b_0\,k\ln 2 + O(1), \qquad b_0 = \frac{11N_c}{48\pi^2},$$

so that β_k decreases *along the blocking and reaches the terminal coupling γ after $n_{\max} = \frac{1}{2b_0\ln 2}\big(\frac{1}{g^2} - \frac{1}{\gamma}\big) + O(1)$ steps, giving $2^{n_{\max}} = e^{C/g^2 + O(1)}$ with $C = \ln 2/(2b_0) = 24\pi^2\ln 2/(11N_c)$.*

Remark 3.2 (Erratum in v1, corrected in v2)*.* v1 stated (c) as $\beta_k = \beta + 2b_0 k\ln 2$: the sign is inverted (the coupling must *grow* toward the infrared, i.e. β_k decreases, or no terminal coupling $\gamma = O(1)$ is ever reached from $\beta \gg 1$), and the normalization conflates β with $1/g^2$ (missing factor $2N_c$). The corrected flow gives exactly the same $n_{\max}$ and $C(N_c) = 24\pi^2\ln 2/(11N_c)$ that v1 used in Section 8, so no downstream constant changes; the v1 formula was a typo inconsistent with v1's own use of it.

4. Ricci curvature of the orbit space

Theorem 4.1 (Orbit space Ricci bound)**.** *The orbit space $\mathcal{B} = \mathcal{A}/\mathcal{G}$ satisfies $\mathrm{Ric}_{\mathcal{B}} \geq N_c/4$ on the smooth stratum $\mathcal{B}^*$.*

Proof. Step 1: Ricci curvature of $\mathcal{A}$. Each factor $\mathrm{SU}(N_c)$ with bi-invariant metric $\langle X, Y\rangle = -\operatorname{tr}(XY)$ has sectional curvature $K(X,Y) = \frac{1}{4}\|[X,Y]\|^2/(\|X\|^2\|Y\|^2 - \langle X,Y\rangle^2)$ and Ricci curvature $\mathrm{Ric}_{\mathrm{SU}(N_c)}(X,X) = \frac{1}{4}\sum_i \|[X,e_i]\|^2 = \frac{N_c}{2}\|X\|^2$ (the $-B/4$ formula for the Killing form $B = 2N_c\operatorname{tr}$; see Remark 4.3). For the product $\mathcal{A} = \prod_\ell \mathrm{SU}(N_c)$: $\mathrm{Ric}_{\mathcal{A}} = \bigoplus_\ell \mathrm{Ric}_{\mathrm{SU}(N_c)}$, so $\mathrm{Ric}_{\mathcal{A}} \geq N_c/4$ (indeed $\geq N_c/2$).

Step 2: O'Neill's formula. On $\mathcal{B}^*$: the gauge group acts freely and $\pi : \pi^{-1}(\mathcal{B}^*) \to \mathcal{B}^*$ is a Riemannian submersion (Proposition 2.2). O'Neill's formula [7] gives:

$$\mathrm{Ric}_{\mathcal{B}^*}(\pi_* X, \pi_* X) = \mathrm{Ric}_{\mathcal{A}}(X,X) + 3\sum_i \|A_X E_i\|^2 \;\geq\; \mathrm{Ric}_{\mathcal{A}}(X,X) \;\geq\; \frac{N_c}{4}\|X\|^2. \qquad \square$$

Remark 4.2 (Treatment of orbifold singularities)*.* The Ricci bound $\mathrm{Ric}_{\mathcal{B}} \geq N_c/4$ is proved on $\mathcal{B}^*$, which has full measure in $\mathcal{B}$ (Proposition 2.2: the singular stratum has codimension ≥ 2). For the Bakry–Émery criterion and the Holley–Stroock lemma: the log-Sobolev inequality

$$\int f^2 \ln f^2\, d\mu - \int f^2 d\mu\, \ln \int f^2 d\mu \;\leq\; \frac{2}{\alpha}\int |\nabla f|^2\, d\mu$$

involves only the gradient ∇f, which is well-defined μ-a.e. (since $\mathcal{B}_{\mathrm{sing}}$ has μ-measure zero for any Gibbs measure μ absolutely continuous with respect to the Riemannian volume). The curvature-dimension condition $CD(N_c/4, \infty)$ holds on $\mathcal{B}^*$ by Theorem 4.1, hence on all of $\mathcal{B}$ in the μ-a.e. sense required for the Bakry–Émery theorem. This is the standard treatment of log-Sobolev inequalities on Riemannian orbifolds; see [9] for spectral theory on orbifolds and [10] for curvature-dimension conditions on metric measure spaces (which subsume orbifolds).

Remark 4.3 (New in v2: sharp constant, and status of this section). (i) v1 stated $\mathrm{Ric}_{\mathrm{SU}(N_c)}(X,X) = \frac{N_c}{4}\|X\|^2$, citing [8, Prop. 3.18]. With the metric $\langle X,Y\rangle = -\operatorname{tr}(XY)$ the sharp value is $\frac{N_c}{2}\|X\|^2$: for bi-invariant metrics $\mathrm{Ric} = -\frac{1}{4}B$, and $B(X,Y) = 2N_c \operatorname{tr}(XY) = -2N_c\langle X,Y\rangle$ for $\mathfrak{su}(N_c)$. The verification suite confirms $\mathrm{Ric}(X,X)/\|X\|^2 = 1.000$ for SU(2) and 1.500 for SU(3) via the structure-constant formula. The v1 constant is thus an undercount by a factor of 2; the stated bound $\mathrm{Ric}_{\mathcal{B}} \geq N_c/4$ remains true (just not sharp), so this erratum is in the *favorable* direction. (ii) We record that Theorem 4.1 is not used in the proof chain of Theorem 1.1: the Witten-Laplacian route of Section 6 bypasses the Bakry–Émery/Holley–Stroock argument entirely. The section is retained as geometric context.

5. Morse–Bott structure of the Wilson potential

Theorem 5.1 (Morse–Bott structure on the smooth stratum). *The Wilson potential $V_W : \mathcal{B} \to \mathbb{R}$ is Morse–Bott along the smooth stratum of its critical set (see Proposition 5.3 for the failure at the orbifold locus; in v1 this theorem asserted the Morse–Bott property without restriction):*

(a) *The critical set is $\mathrm{Crit}(V_W) = \mathcal{M}_{\mathrm{flat}} \cong T^r/W$, the moduli space of flat connections, where $r = N_c - 1$ and $W = S_{N_c}$ is the Weyl group.*

(b) *The normal Hessian $\mathrm{Hess}^{\perp}(V_W)|_{\mathcal{M}_{\mathrm{flat}}}$ has eigenvalues*

$$\nu_{k,a}(\theta) = 4\sum_{\mu=1}^{d-1} \sin^2\Big(\frac{\pi k_\mu + \alpha_a\cdot\theta_\mu/2}{L}\Big) \tag{2}$$

for spatial momentum $k \neq 0$ and root α_a. At $\theta = 0$: $\nu_{k,a}(0) = 4\sum_\mu \sin^2(\pi k_\mu/L) \geq \nu_{\min} := 4\sin^2(\pi/L) > 0$.

Proof. (a) $V_W([U]) = 0$ iff every plaquette is the identity, iff U is a flat connection. On the torus, flat connections modulo gauge are classified by their commuting holonomies, giving $\mathcal{M}_{\mathrm{flat}} \cong \mathrm{Hom}(\mathbb{Z}^d, T)/W \cong T^r/W$.

(b) Expand V_W to second order around a flat connection $U_\mu^{(0)}(x) = \exp(i\theta_\mu\cdot H/L)$ (where H is a basis of the Cartan subalgebra) in the direction of a gauge-invariant fluctuation $\eta_\mu^a(x) \sim e^{2\pi i k\cdot x/L}$ in the adjoint representation. The covariant finite-difference Hessian gives (2) by direct computation (see e.g. [13] for the continuum analogue). The verification suite validates (2) at $\theta = 0$ non-perturbatively: the finite-difference Hessian of S_W on the full SU(2) configuration space of a 2^3 spatial lattice agrees entry-by-entry with $(1/N_c)(d^T d)\otimes I_{N_c^2-1}$ (lattice Maxwell per adjoint color), with nonzero spectrum $\{\hat{k}^2/N_c\}$. (The overall factor $1/N_c$ relative to (2) is the metric normalization $\langle X,Y\rangle = -\operatorname{tr}(XY)$; it is absorbed into the constants c_1, C' and does not affect positivity.) □

Lemma 5.2 (Born–Oppenheimer potential minimum). *Define the Born–Oppenheimer potential on $\mathcal{M}_{\mathrm{flat}}$ by*

$$V_{\mathrm{BO}}(\theta) = \frac{1}{2}\sum_{k\neq 0}\sum_a \Big[\,\omega_{k,a}(\theta) - \omega_{k,a}(0)\,\Big], \qquad \omega_{k,a}(\theta)^2 := \hat{k}^2 + m_a^2(\theta), \tag{3}$$

where $\hat{k}^2 = 4\sum_\mu \sin^2(\pi k_\mu/L)$ and $m_a^2(\theta) = 4\sum_\mu \sin^2(\alpha_a\cdot\theta_\mu/2)$. (In v1, (3) was written with $\omega_{k,a} = \sqrt{\nu_{k,a}}$ from (2); see Remark 5.4 — that version is not positive and is corrected here to the dispersion actually used in the v1 proof.) Then:

(i) *$V_{\mathrm{BO}}(\theta) \geq 0$ with equality iff $\theta = 0$ in T^r/W.*

(ii) *The Hessian of V_{BO} at $\theta = 0$ is positive definite:*

$$\frac{\partial^2 V_{\mathrm{BO}}}{\partial\theta_{i,\mu}\partial\theta_{j,\nu}}\Big|_{\theta=0} = \delta_{\mu\nu}\, N_c\, S_1(L)\, \delta_{ij} \tag{4}$$

where $S_1(L) = \sum_{k\neq 0}\big(4\sum_\mu \sin^2(\pi k_\mu/L)\big)^{-1/2} > 0$ is a finite positive constant depending on the lattice size. (v1 stated $\frac{N_c}{2}S_1$; the correct constant is $N_c S_1$, Remark 5.4.)

Proof. (i) For each fixed spatial momentum $k \neq 0$, define $F_k(\theta) = \sum_a\big[\sqrt{\hat{k}^2 + m_a^2(\theta)} - |\hat{k}|\big]$. Since $t \mapsto \sqrt{\hat{k}^2 + t} - |\hat{k}|$ is non-negative and strictly increasing for $t \geq 0$: each summand in F_k is non-negative, and $F_k(\theta) \geq 0$. Moreover, $F_k(\theta) = 0$ iff $m_a^2(\theta) = 0$ for every root α_a. Since the roots of $\mathrm{SU}(N_c)$ span $\mathfrak{t}^*$ for $N_c \geq 2$: $m_a^2(\theta) = 0$ for all a implies $\theta_\mu = 0$ for all μ (modulo the coroot lattice, i.e., $\theta = 0$ in T^r/W). Therefore $V_{\mathrm{BO}}(\theta) = \frac{1}{2}\sum_{k\neq 0} F_k(\theta) \geq 0$, with equality if and only if $\theta = 0$.

(ii) Computing the Hessian at $\theta = 0$:

$$\frac{\partial^2 V_{\mathrm{BO}}}{\partial\theta_{i,\mu}\partial\theta_{j,\nu}}\Big|_0 = \frac{1}{2}\sum_{k\neq 0}\sum_a \frac{1}{2|\hat{k}|}\frac{\partial^2 m_a^2}{\partial\theta_{i,\mu}\partial\theta_{j,\nu}}\Big|_0,$$

using $\partial_\theta m_a^2|_0 = 0$ so only the second derivative of m_a^2 contributes. Now $m_a^2(\theta) = 4\sum_\rho \sin^2(\alpha_a \cdot \theta_\rho/2)$, so

$$\frac{\partial^2 m_a^2}{\partial\theta_{i,\mu}\partial\theta_{j,\nu}}\Big|_0 = 2\,\delta_{\mu\nu}\,(\alpha_a)_i(\alpha_a)_j$$

(v1 dropped the factor 2: $\partial_\varphi^2\, 4\sin^2(\varphi/2)|_0 = 2$). Using $\sum_a(\alpha_a)_i(\alpha_a)_j = 2N_c\delta_{ij}$ (the standard root system identity for $\mathrm{SU}(N_c)$, verified for $N_c \leq 5$ in the suite):

$$\frac{\partial^2 V_{\mathrm{BO}}}{\partial\theta_{i,\mu}\partial\theta_{j,\nu}}\Big|_0 = \delta_{\mu\nu}\delta_{ij}\cdot N_c\sum_{k\neq 0}\frac{1}{|\hat{k}|} = \delta_{\mu\nu}\delta_{ij}\cdot N_c\, S_1(L).$$

Since $S_1(L) > 0$ for any $L \geq 2$: the Hessian is $(N_c S_1)\cdot I > 0$. The finite-difference Hessian in the suite reproduces $N_c S_1(L)\, I$ to relative precision 10^{-6} for $(N_c, L) \in \{(2,4),(2,8),(3,4)\}$. ☐

Proposition 5.3 (New in v2: Morse–Bott failure at the orbifold locus)**.** *At points $\theta \in \mathcal{M}_{\mathrm{flat}}$ with enhanced stabilizer — the orbifold locus of T^r/W, which includes $\theta = 0$ — the Morse–Bott non-degeneracy fails: in addition to the tangent directions of the critical set (gauge orbit plus Cartan moduli), the Hessian of S_W has $(d-1)(N_c^2 - 1 - r)$ zero modes from the $k = 0$ non-Cartan constant fluctuations, along which S_W is* quartic*. At generic θ (trivial enhancement) these modes acquire mass $m_a^2(\theta) > 0$ and the Morse–Bott property is restored.*

Proof. The $k = 0$ adjoint constant modes have normal frequency $m_a(\theta)$, which vanishes precisely when $\alpha_a \cdot \theta = 0 \bmod 2\pi$, i.e. at enhanced-stabilizer points. At $\theta = 0$: constant non-commuting fluctuations X_μ give $U_\square = \exp([X_\mu, X_\nu] + \cdots)$, so $S_W \sim \frac{1}{4N_c}\sum \|[X_\mu, X_\nu]\|^2$ is quartic — these are the toron modes of [15, 16]. Numerically (suite, SU(2), spatial lattice 2^3, finite-difference Hessian on all 72 configuration coordinates, no linearization): at $\theta = 0$ the kernel has dimension $30 = 21$ (gauge orbit) $+\, 3$ (Cartan moduli) $+\, 6$ (non-Cartan $k = 0$), exceeding $\dim T\widetilde{\mathcal{M}}_{\mathrm{flat}} = 24$; the six extra directions satisfy $f(2t)/f(t) = 15.99 \approx 2^4$ (quartic); a *single* constant non-Cartan mode is exactly flat (it is gauge-conjugate to a Cartan modulus). At generic θ the kernel drops to $26 = 23$ (gauge orbit, stabilizer $= T$)$+3$ (moduli) with strictly positive remaining spectrum. ☐

Remark 5.4 (New in v2: erratum in the v1 Born–Oppenheimer potential)*.* v1 defined V_{BO} by (3) with $\omega_{k,a} = \sqrt{\nu_{k,a}}$ from the twisted-momentum formula (2), but *proved* Lemma 5.2 for the dispersion $\omega_{k,a}^2 = \hat{k}^2 + m_a^2(\theta)$ — two different functions at finite L. The literal v1 definition is not positive: numerically (suite), its Hessian at $\theta = 0$ is $\approx -0.117\,I$ (SU(2), $L = 4$) and the potential reaches -2.45 on the scan domain, so v1's Lemma 5.2(i)–(ii) are *false* for the object v1 defined (the excluded $k = 0$ term, which carries the stabilizing mass kink, is precisely what Proposition 5.3 is about). v2 therefore takes (3) with $\omega^2 = \hat{k}^2 + m_a^2$ — the dispersion the v1 proof actually used — as the definition; for it, positivity holds (verified globally on a 15^3 grid over the coroot cell) and the Hessian is (4) with the corrected constant $N_c S_1$ (v1's $\frac{N_c}{2} S_1$ dropped a factor 2 in $\partial^2 m_a^2$). Consequently $\omega_{\min} = \sqrt{N_c S_1}$, a factor $\sqrt{2}$ *larger* than v1's value: the correction is favorable and propagates to (5), (6) and Theorem 1.1.

6. Witten Laplacian spectral gap at the terminal scale

Theorem 6.1 (Semiclassical gap, conditional on (H-MB) and (H-HAM))**.** *Let $\mathcal{B}_{n_{\max}}$ be the orbit space at Balaban's terminal scale with $L_{n_{\max}} \le C_0$. Assume* (H-HAM) *and* (H-MB)*. The spectral gap of the Hamiltonian $H = (\gamma/2)(-\Delta_{\mathcal{B}}) + V_{\text{eff}}$ satisfies*

$$m_{n_{\max}} \;\ge\; \frac{c_1 N_c^{3/2}\,\omega_{\min}}{\gamma^2} \;>\; 0 \tag{5}$$

where $\omega_{\min} = \sqrt{N_c S_1} > 0$ (Lemma 5.2), $\gamma > 0$ is Balaban's terminal coupling and $c_1 > 0$ depends on C_0.

Proof. Step 1: Witten Laplacian formulation. The ground state $\psi_0 = e^{-V_{\text{eff}}/\gamma}/\sqrt{Z}$ gives $m = (\gamma/2)\lambda_1(\Delta_f^{(0)})$ where $f = V_{\text{eff}}/\gamma$ and $\Delta_f^{(0)} = -\Delta + |\nabla f|^2 - \Delta f$.

Step 2: Semiclassical identification. Write $f = \alpha f_0$ with $f_0 = V_W + O(e^{-\kappa})$ and $\alpha = \beta_{n_{\max}}/\gamma = N_c/\gamma^2$. Then $\Delta_f^{(0)} = 4\alpha^2 \Delta_{h,f_0}^{(0)}$ with $h = 1/(2\alpha) = \gamma^2/(2N_c) \le h_0 \ll 1$ (this is where (H-CONST) enters: $h \le h_0$ requires $\gamma \le \gamma_{\max}$). □

Lemma 6.2 (Morse–Bott stability under small perturbation)**.** *Let M be a compact Riemannian manifold and $V : M \to \mathbb{R}$ a Morse–Bott function with critical manifold Σ and normal Hessian eigenvalues $\ge \nu_{\min} > 0$. For any $\delta V \in C^2(M)$ with $\|\delta V\|_{C^2} \le \epsilon$: if $\epsilon < \nu_{\min}/(2C_M)$ (where C_M depends only on the geometry of M and Σ), then $f_0 = V + \delta V$ is also Morse–Bott with:*

(i) *a critical manifold Σ_ϵ that is $O(\epsilon/\nu_{\min})$-close to Σ in C^1;*
(ii) *normal Hessian eigenvalues $\ge \nu_{\min} - C_M\epsilon > \nu_{\min}/2 > 0$.*

Proof. The gradient ∇V vanishes on Σ and the normal Hessian $\operatorname{Hess}^\perp V|_\Sigma$ has smallest eigenvalue $\nu_{\min} > 0$. By the implicit function theorem applied to $\nabla^\perp(V + \delta V) = 0$ in a tubular neighborhood of Σ: for $\|\delta V\|_{C^2} \le \epsilon$ sufficiently small, the zero set Σ_ϵ of $\nabla^\perp f_0$ is a smooth submanifold $O(\epsilon/\nu_{\min})$-close to Σ. The normal Hessian of f_0 at Σ_ϵ satisfies $\operatorname{Hess}^\perp(f_0)|_{\Sigma_\epsilon} = \operatorname{Hess}^\perp(V)|_\Sigma + O(\epsilon)$, giving eigenvalues $\ge \nu_{\min} - C_M\epsilon$. For Balaban's corrections: $\epsilon = O(e^{-\kappa}) \ll \nu_{\min}$ (Theorem 3.1(b)), so both conditions hold. (Caveat added in v2: this lemma applies on the smooth stratum, where $\nu_{\min} > 0$; on the orbifold locus of Proposition 5.3 its hypothesis fails and (H-MB) is invoked instead.) □

Step 3: Lift to the smooth manifold $\mathcal{A}$. The orbit space $\mathcal{B}$ is an orbifold with singularities (Proposition 2.2), which complicates the direct application of Helffer–Sjöstrand theory. We bypass this by lifting the spectral problem to the smooth compact manifold $\mathcal{A}$. Since $f_0 = V_W + O(e^{-\kappa})$ is $\mathcal{G}$-invariant, the Witten Laplacian $\Delta_{h,f_0}^{(0)}$ on $\mathcal{A}$ commutes with the $\mathcal{G}$-action, hence preserves $L^2(\mathcal{A})^{\mathcal{G}}$, and its restriction there is unitarily equivalent (via π^*) to the operator on $\mathcal{B}$. The v1 checklist of the hypotheses of [12, Theorem 1.4] on $\mathcal{A}$ must be amended as follows (v2):

(H1) $\mathcal{A}$ is a compact smooth Riemannian manifold without boundary. ✓

(H2) f_0 is Morse–Bott on $\mathcal{A}$: *holds only on the smooth stratum of* $\widetilde{\mathcal{M}}_{\text{flat}} = \pi^{-1}(\mathcal{M}_{\text{flat}})$. At enhanced-stabilizer points — including the fiber over $\theta = 0$ — the normal Hessian has the additional quartic zero modes of Proposition 5.3, and the Morse–Bott property *fails*. v1 asserted that the normal space consists exclusively of $k \neq 0$ fluctuations; this omitted the $k = 0$ non-Cartan modes, which are normal to $\widetilde{\mathcal{M}}_{\text{flat}}$ (they leave the flat set at quartic order) yet have vanishing Hessian. Here (H-MB) is invoked.

(H3) f_0 attains its global minimum on $\widetilde{\mathcal{M}}_{\text{flat}}$: $V_W = 0$ on $\widetilde{\mathcal{M}}_{\text{flat}}$ and $V_W > 0$ elsewhere. ✓

(H4) Non-degeneracy of the normal Hessian in the gauge-invariant sector: holds for the $k \neq 0$ sector ($\nu_{k,a}(0) \geq \nu_{\min} > 0$, Theorem 5.1(b)); fails for the $k = 0$ non-Cartan sector at the orbifold locus. (H-MB).

Remark 6.3. By lifting to $\mathcal{A}$, we apply the Helffer–Sjöstrand theorem on a smooth compact manifold, avoiding any need to extend semiclassical spectral theory to orbifolds. The gauge invariance of f_0 ensures that the restriction to $L^2(\mathcal{A})^{\mathcal{G}}$ is spectrally equivalent to the operator on $\mathcal{B}$. The critical manifold $\widetilde{\mathcal{M}}_{\text{flat}} \subset \mathcal{A}$ is smooth (a union of smooth $\mathcal{G}$-orbits), even though its image $\mathcal{M}_{\text{flat}} \subset \mathcal{B}$ is an orbifold. (v2: note that the lift removes the *orbit-space* singularities but not the failure of normal non-degeneracy of Proposition 5.3, which lives already on $\mathcal{A}$: the two are distinct phenomena that happen to sit over the same locus.)

Lemma 6.4 (Equivariant spectral restriction)**.** *Let M be a compact Riemannian manifold, G a compact Lie group acting on M by isometries, and $f : M \to \mathbb{R}$ a G-invariant smooth function. Then:*

(i) *The Witten Laplacian $\Delta^{(0)}_{h,f}$ on $L^2(M)$ commutes with the G-action and preserves each isotypic component $L^2(M)^\rho$.*

(ii) *The spectral gap of $\Delta^{(0)}_{h,f}|_{L^2(M)^G}$ satisfies $\lambda_1^{G\text{-}inv} \geq \lambda_1^{\text{full}}$. In particular, if the Helffer–Sjöstrand asymptotics give $\lambda_1^{\text{full}} \geq c\sqrt{h}$, then $\lambda_1^{G\text{-}inv} \geq c\sqrt{h}$.*

Proof. (i) Since f is G-invariant: ∇f is G-equivariant and Δf is G-invariant. The three terms of $\Delta^{(0)}_{h,f} = -h^2\Delta + |\nabla f|^2 - h\Delta f$ are each G-equivariant, so their sum commutes with the G-action, preserving the isotypic decomposition.

(ii) The ground state $\psi_0 = e^{-f/h}/\|e^{-f/h}\|$ is G-invariant, so $\lambda_0 = 0$ lies in the G-invariant sector. By the minimax principle:

$$\lambda_1^{G\text{-inv}} = \min_{\substack{\varphi \in L^2(M)^G \\ \varphi \perp \psi_0,\ \|\varphi\|=1}} \langle \varphi, \Delta^{(0)}_{h,f}\varphi \rangle \ \geq \ \min_{\substack{\varphi \in L^2(M) \\ \varphi \perp \psi_0,\ \|\varphi\|=1}} \langle \varphi, \Delta^{(0)}_{h,f}\varphi \rangle = \lambda_1^{\text{full}}. \qquad \square$$

Lemma 6.5 (Equivariant reduction of the effective operator)**.** *Since $\Delta^{(0)}_{h,f_0}$ commutes with $\mathcal{G}$ (Lemma 6.4(i)), the Helffer–Sjöstrand parametrix construction* [12, Section 4] *preserves $\mathcal{G}$-equivariance at every order in h. In particular:*

(i) *The effective operator H_{eff} on $\widetilde{\mathcal{M}}_{\text{flat}}$ and all error terms $O(h^{3/2})$ in the asymptotic expansion are $\mathcal{G}$-equivariant.*

(ii) *The spectral gap of $\Delta^{(0)}_{h,f_0}|_{L^2(\mathcal{A})^{\mathcal{G}}}$ in the low-lying cluster $[0, Ch]$ equals the spectral gap of $H_{\text{eff}}|_{L^2(\widetilde{\mathcal{M}}_{\text{flat}})^{\mathcal{G}}}$.*

(iii) *Since $\widetilde{\mathcal{M}}_{\text{flat}} = \mathcal{G} \cdot \mathcal{M}_{\text{flat}}$: gauge-invariant functions on $\widetilde{\mathcal{M}}_{\text{flat}}$ correspond to functions on $\mathcal{M}_{\text{flat}} = T^r/W$, and $H_{\text{eff}}|_{L^2(\widetilde{\mathcal{M}}_{\text{flat}})^{\mathcal{G}}}$ is unitarily equivalent to $H^W_{\text{eff}} = -h^2\Delta_{T^r} + hV_{\text{BO}} + O(h^{3/2})$ acting on W-invariant functions on T^r.*

(v2: statements (i)–(iii) are as in v1 and their proofs below are unchanged, but they presuppose the applicability of [12, Section 4], i.e. (H2)/(H4) — so on the quartic locus they too are conditional on (H-MB).)

Proof. (i) The parametrix for $\Delta^{(0)}_{h,f_0}$ near $\widetilde{\mathcal{M}}_{\text{flat}}$ is constructed via a formal power series in h: normal-form coordinates, Taylor expansion of the symbol, and inversion order by order [12, Section 4]. Each step involves only algebraic operations on the symbol of $\Delta^{(0)}_{h,f_0}$ and the geometry of $\widetilde{\mathcal{M}}_{\text{flat}} \subset \mathcal{A}$. Since $\Delta^{(0)}_{h,f_0}$ is $\mathcal{G}$-equivariant, its symbol is $\mathcal{G}$-invariant; since $\widetilde{\mathcal{M}}_{\text{flat}}$ is $\mathcal{G}$-invariant, the normal-form coordinates can be chosen $\mathcal{G}$-equivariantly (using the $\mathcal{G}$-equivariant tubular neighborhood from the slice theorem). Therefore each term in the parametrix expansion, including all remainders $O(h^{3/2})$, is $\mathcal{G}$-equivariant.

(ii) Since the projection $\Pi_{[0,Ch]}$ onto the low-lying spectral subspace commutes with $\mathcal{G}$, the low-lying spectrum restricted to $L^2(\mathcal{A})^{\mathcal{G}}$ is the intersection of the low-lying spectrum with the $\mathcal{G}$-invariant sector. The effective operator H_{eff} represents this restricted spectrum.

(iii) The map $\varphi \mapsto \varphi|_{\mathcal{M}_{\text{flat}}}$ is a unitary equivalence from $L^2(\widetilde{\mathcal{M}}_{\text{flat}})^{\mathcal{G}}$ to $L^2(\mathcal{M}_{\text{flat}})^W = L^2(T^r)^W$ (since $\mathcal{G}$ acts transitively on each fiber of $\widetilde{\mathcal{M}}_{\text{flat}} \to \mathcal{M}_{\text{flat}}$ with finite stabilizer W). Under this equivalence, H_{eff} becomes $H^W_{\text{eff}} = -h^2\Delta_{T^r} + hV_{\text{BO}} + O(h^{3/2})$ on $L^2(T^r)^W$. □

Step 4: Application of the theorem. Under (H-MB) — which supplies what [12, Theorem 1.4] would supply if its hypotheses held on all of $\widetilde{\mathcal{M}}_{\text{flat}}$ — the low-lying spectrum of $\Delta^{(0)}_{h,f_0}$ in $[0, Ch]$ is determined by an effective operator on $\widetilde{\mathcal{M}}_{\text{flat}}$, and the gap to the next spectral cluster is $\geq c_2 h$. By Lemma 6.5: the restriction to the gauge-invariant sector gives $H^W_{\text{eff}} = -h^2\Delta_{T^r} + hV_{\text{BO}} + O(h^{3/2})$ on $L^2(T^r)^W$. Since W is a finite group, restriction to W-invariant functions only increases the spectral gap (Lemma 6.4 with $G = W$).

Step 5: Effective Hamiltonian spectral gap. Near $\theta = 0$: $V_{\text{BO}}(\theta) \approx \frac{1}{2}\omega^2_{\min}|\theta|^2$ with $\omega^2_{\min} = N_c S_1$ (Lemma 5.2, corrected constant). The leading eigenvalues of H_{eff} near $\theta = 0$ are those of the harmonic oscillator $-h^2\Delta + \frac{h\omega^2_{\min}}{2}|\theta|^2$, giving spectral gap $\lambda^{\text{eff}}_1 = \sqrt{h}\,\omega_{\min}(1 + O(h^{1/2}))$. (v2: it is exactly this harmonic approximation that fails in the quartic directions of Proposition 5.3; (H-MB) asserts the conclusion regardless.)

Step 6: Conversion to mass gap.

(6)
$$m_{n_{\max}} = \frac{\gamma}{2}\lambda_1(\Delta^{(0)}_f) = \frac{\gamma}{2}\cdot 4\alpha^2 \cdot \lambda^{\text{eff}}_1 = \frac{\gamma}{2}\cdot\frac{4N_c^2}{\gamma^4}\cdot\sqrt{\frac{\gamma^2}{2N_c}}\,\omega_{\min}(1+O(\gamma)) = \frac{\sqrt{2}\,N_c^{3/2}\,\omega_{\min}}{\gamma^2}(1+O(\gamma)).$$

For $\gamma \leq \gamma_0$ (sufficiently small that the $O(\gamma)$ correction is bounded by $1/2$):

$$m_{n_{\max}} \;\geq\; \frac{N_c^{3/2}\omega_{\min}}{\sqrt{2}\,\gamma^2} =: \frac{c_1 N_c^{3/2}\omega_{\min}}{\gamma^2}, \qquad c_1 = 1/\sqrt{2}.$$

Remark 6.6. The bound (6) gives $m_{n_{\max}} = O(N_c^{3/2}/\gamma^2)$. Since γ is Balaban's fixed constant, $m_{n_{\max}}$ is a fixed positive number. That $m_{n_{\max}} \gg 1$ in terminal lattice units means the correlation length $\xi = 1/m_{n_{\max}} \ll a_{n_{\max}}$ (sub-lattice). This does not by itself invalidate the semiclassical analysis: the Helffer–Sjöstrand asymptotics are controlled by the semiclassical parameter $h = \gamma^2/(2N_c) \ll 1$, not by the resulting gap. Physically, $m_{n_{\max}} \gg 1$ is consistent with the terminal-scale theory describing degrees of freedom tightly confined near $\mathcal{M}_{\text{flat}}$: Balaban's RG has integrated out all fluctuations above $a_{n_{\max}}$, and the remaining effective theory is strongly confining in lattice units while its physical mass gap $m_{\text{phys}} = m_{n_{\max}}/a_{n_{\max}}$ remains $O(\Lambda_{\text{QCD}})$.

(v2: note however that $m_{n_{\max}} \gg 1$ is precisely what collides with $\kappa = O(1)$ in Theorem 7.2; see Remark 7.3 and (H-CONST).)

Remark 6.7 (New in v2: the quartic sector). The $k = 0$ non-Cartan modes at $\theta = 0$ (Proposition 5.3) are not covered by Steps 4–5: for them the normal expansion of the Witten Laplacian is governed by a quartic, not quadratic, potential, and the LPNV cluster theorem does not apply. The small-volume literature [15, 16] constructs precisely the effective Hamiltonian for this sector (for SU(2) and SU(3) on T^3) and finds a discrete low-lying spectrum with gaps of order $g^{2/3}$ in appropriate units — strictly positive, which is what (H-MB) needs, but with different h-scaling than the harmonic $\sqrt{h}\,\omega_{\min}$. A proof of (H-MB) would need to merge that analysis with the LPNV parametrix away from the orbifold locus; we do not attempt this here.

7. Transfer matrix scaling with controlled errors

Lemma 7.1 (Approximate temporal factorization)**.** *For Balaban's effective action S_k on Λ_k with $L_k \le C_0$: decompose $S_k = S_k^{\rm nn} + S_k^{\rm lr}$ where $S_k^{\rm nn}$ couples only adjacent temporal slices. Let T_k be the transfer matrix defined by $S_k^{\rm nn}$. Then:*

$$\left|Z_k - \operatorname{Tr}(T_k^{L_k})\right| \;\le\; \operatorname{Tr}(T_k^{L_k}) \cdot \delta$$

with $\delta = O(e^{-\kappa})$, where κ is Balaban's universal decay constant.

Proof. The long-range part satisfies $|S_k^{\rm lr}| \le C \cdot L_k^4 \cdot e^{-2\kappa}$ (only terms with temporal extent $\ge 2a_k$ contribute, each bounded by $e^{-2\kappa}$, with $O(L_k^4)$ such terms). Then $Z_k = \operatorname{Tr}(T_k^{L_k})(1+\delta)$ with $|\delta| \le e^{CL_k^4 e^{-2\kappa}} - 1 = O(e^{-\kappa})$ for C_0 fixed. □

Theorem 7.2 (Mass gap scaling; statement corrected in v2)**.** *For lattice sizes $M = L_k$ in the finite window $C \le M \le \kappa/(2m_k)$:*

$$m_{k+1} = 2m_k + O\big(M^{-1} e^{-(\kappa - m_k M)}\big).$$

If moreover $m_k L_k \le \kappa/2$ for all $k \le n_{\max}$ (part of (H-CONST)*), then $m_0 = m_{n_{\max}}/2^{n_{\max}} + O(e^{-\kappa/2})$. (v1 stated "$m_{k+1} = 2m_k + O(e^{-(\kappa - m_k)M})$ for $M = L_k \to \infty$"; see Remark 7.4.)*

Proof. Step 1: Approximate trace identity. By $Z_k = Z_{k+1}$ (exact, Theorem 3.1(a)) and Lemma 7.1:

$$\operatorname{Tr}(T_{k+1}^{M/2})(1+\delta_{k+1}) = \operatorname{Tr}(T_k^M)(1+\delta_k) \tag{7}$$

with $|\delta_k| = O(e^{-\kappa})$, for all admissible $M = L_k$ divisible by $2^{n_{\max}-k}$.

Step 2: Spectral expansion. For M in the window ($1 \ll M$):

$$\operatorname{Tr}(T_k^M) = (\lambda_0^{(k)})^M\big[1 + e^{-m_k M} + O(e^{-m_k' M})\big]$$

where $m_k = -\ln(\lambda_1^{(k)}/\lambda_0^{(k)})$ and $m_k' = -\ln(\lambda_2^{(k)}/\lambda_0^{(k)}) > m_k$.

Step 3: Leading order. Substituting, taking logs and dividing by M: $\lambda_0^{(k+1)} = (\lambda_0^{(k)})^2(1 + O(e^{-\kappa}/M))$.

Step 4: Subleading order (corrected in v2). Dividing by $(\lambda_0^{(k)})^M$:

$$1 + e^{-m_{k+1}M/2} + O(e^{-m_{k+1}'M/2}) = \big[1 + e^{-m_k M} + O(e^{-m_k' M})\big] \cdot \big[1 + O(e^{-\kappa})\big].$$

The signal $e^{-m_k M}$ dominates the error $O(e^{-\kappa})$ if and only if $m_k M < \kappa$, i.e. $M < \kappa/m_k$ (v1 printed the reverse inequality). On the window $M \le \kappa/(2m_k)$:

$$e^{-m_{k+1}M/2} = e^{-m_k M}\big(1 + O(e^{-(\kappa - m_k M)}) + O(e^{-(m_k' - m_k)M})\big),$$

and taking logs and dividing by $M/2$:

$$m_{k+1} = 2m_k + O\big(M^{-1} e^{-(\kappa - m_k M)}\big) + O\big(M^{-1} e^{-(m_k' - m_k)M}\big).$$

No $M \to \infty$ limit is available (or needed): M is capped by $\kappa/(2m_k)$, and the error is controlled on the finite window. Iterating downward with $m_k = m_{n_{\max}}/2^{n_{\max}-k}$ (so the window constraint is most restrictive at $k = n_{\max}$, where it reads $m_{n_{\max}} L_{n_{\max}} \le \kappa/2$) accumulates a total error $O(e^{-\kappa/2})$, giving $m_0 = m_{n_{\max}}/2^{n_{\max}} + O(e^{-\kappa/2})$. □

Remark 7.3 (Verification of $\kappa \gg m_k$; revised in v2). The scaling argument requires $m_k L_k \le \kappa/2$ for all $k \le n_{\max}$; the most restrictive case is $k = n_{\max}$. From Theorem 6.1, the semiclassical analysis also gives a matching upper bound $m_{n_{\max}} \le C' N_c^{3/2}/\gamma^2$ for an explicit constant C', so the condition becomes $\kappa > C' N_c^{3/2}/\gamma^2 \cdot 2L_{n_{\max}}/\ldots$ — up to the $O(1)$ factor $2L_{n_{\max}} \le 2C_0$, the requirement is

$$\kappa \;\gtrsim\; \frac{C' N_c^{3/2}}{\gamma^2}.$$

v1 argued that this "constrains γ from below: $\gamma > \gamma_{\min} = (C' N_c^{3/2}/\kappa)^{1/2}$", imposed the window $\gamma_{\min} < \gamma \le \gamma_0$, and asserted it "satisfiable for N_c fixed and κ sufficiently large (as guaranteed by Balaban's construction)". v2 corrects the logic: κ is a fixed output of Balaban's cluster expansion and cannot be taken large at will, while the semiclassical side (Theorem 6.1) simultaneously requires $\gamma^2 \le 2N_c h_0$ with $h_0 \ll 1$. The two requirements are jointly satisfiable iff $\kappa > C'\sqrt{N_c}/(2h_0)$: numerically (suite), for $C' = 1$, $N_c \in \{2, 3\}$ and $h_0 \le 0.1$ the window is *empty* for every $\kappa \le 5$, and opens only above $\kappa \approx 7$–9. Whether Balaban's constants satisfy this is not established anywhere we know of; this is Hypothesis (H-CONST), and it is the single most load-bearing analytic assumption of the paper.

Remark 7.4 (New in v2: erratum in the v1 extraction regime). v1's Step 4 read: "For $M > \kappa/m_k$ (so that $e^{-m_k M} \gg e^{-\kappa}$)" — the parenthetical requires $m_k M < \kappa$, i.e. exactly the *opposite* regime $M < \kappa/m_k$ — and then took $M \to \infty$, in which limit the signal $e^{-m_k M}$ is buried under the fixed $O(e^{-\kappa})$ error and the extraction fails. The corrected finite-window statement is Theorem 7.2. The failure mode is not hypothetical: in the synthetic transfer-matrix demonstration of the suite ($m_k = 0.3$, $m'_k = 0.5$, $\kappa = 5$), the windowed extraction returns $m_{k+1} = 2m_k$ within its error bound for $8 \le M \le 12$, while at $M = 40 > \kappa/m_k$ the extracted value is off by 58% (error sign $+$) or the log argument is negative (error sign $-$), and as $M \to \infty$ the extracted mass tends to 0, not $2m_k$.

8. Proof of the main theorem

Proof of Theorem 1.1. Assume (H-BAL), (H-CONST), (H-MB), (H-HAM). Choose L with $L_{n_{\max}} = L/2^{n_{\max}} \in [2, C_0]$.

By Theorem 6.1 (under (H-MB), (H-HAM)): $m_{n_{\max}} \ge c_1 N_c^{3/2} \omega_{\min}/\gamma^2 > 0$.

By Theorem 7.2 (under the window condition of (H-CONST)): $m_0 = m_{n_{\max}}/2^{n_{\max}} + O(e^{-\kappa/2})$.

By Theorem 3.1(c): $2^{n_{\max}} = e^{C/g^2 + O(1)}$ with $C = \ln 2/(2b_0) = 24\pi^2 \ln 2/(11N_c)$.

Since $e^{-\kappa/2} \ll m_{n_{\max}}/2^{n_{\max}}$ cannot hold for fixed κ once g is small enough ($m_{n_{\max}}/2^{n_{\max}}$ is exponentially small in $1/g^2$), the additive error must be tracked multiplicatively: iterating Theorem 7.2 multiplicatively, each halving step carries a relative error $1 + O(e^{-(\kappa - m_k L_k)})$, and under $m_k L_k \le \kappa/2$ the product over $k \le n_{\max}$ is $e^{O(n_{\max} e^{-\kappa/2})} = e^{O(1)}$ for $n_{\max} \le e^{\kappa/2}$, which holds in the stated range $L \le C_0 e^{C/g^2}$ provided $g^2 \ge c/\kappa$ suitably; for smaller g the constant degrades explicitly. Hence

$$m_{\mathrm{gap}} = m_0 \;\ge\; \frac{c_1 N_c^{3/2} \omega_{\min}}{2\gamma^2} \cdot e^{-C/g^2 - O(1)} = c(N_c) \cdot e^{-C/g^2}$$

where $c(N_c) = c_1 N_c^{3/2} \omega_{\min}/(2\gamma^2) \cdot e^{-O(1)} > 0$ is a fixed positive constant. □

9. Discussion

9.1. **Summary.** Theorem 1.1 establishes a positive mass gap for $\mathrm{SU}(N_c)$ lattice Yang–Mills in $d = 4$ at weak coupling, *conditionally* on the four hypotheses of Section 1.3. Relative to v1, the honest accounting is: the orbifold-singularity issue in the *orbit space* is genuinely handled by the lift to $\mathcal{A}$ (Remark 6.3), but the Morse–Bott degeneracy at the orbifold locus of the *critical manifold* is not (Proposition 5.3, (H-MB)); the Helffer–Sjöstrand hypotheses are verified only on the smooth stratum; the BO potential minimum survives after correcting its definition (Lemma 5.2); and the constants question (H-CONST) is open.

9.2. **Continuum limit.** The physical mass gap $m_{\mathrm{phys}} = m_0/a_0$ satisfies, using $a_0\Lambda = (b_0 g^2)^{-b_1/(2b_0^2)} e^{-1/(2b_0 g^2)}(1+O(g^2))$:

$$m_{\mathrm{phys}} \;\geq\; c \cdot \Lambda \cdot (b_0 g^2)^{b_1/(2b_0^2)} \cdot e^{(1-\ln 2)/(2b_0 g^2)}.$$

Since $(1 - \ln 2)/(2b_0) > 0$: the exponential diverges as $g^2 \to 0$, but the $(b_0 g^2)^{b_1/(2b_0^2)}$ factor vanishes. For any fixed $g^2 \leq g_0^2$: the bound gives $m_{\mathrm{phys}} \geq c'(N_c, g^2) \cdot \Lambda > 0$. However, we do not establish a uniform positive lower bound as $g^2 \to 0$ ($a \to 0$): this would require controlling the $O(1)$ constants in Balaban's RG with greater precision than is currently available.

9.3. **Limitations.**

(1) $L \leq C_0 e^{C/g^2}$ (finite volume).
(2) Restricted to $d = 4$ (Balaban's results).
(3) No uniform continuum mass gap as $a \to 0$.
(4) The semiclassical bound (Theorem 6.1) uses Helffer–Sjöstrand asymptotics on the smooth manifold $\mathcal{A}$ restricted to gauge-invariant functions; the hypotheses are verified only away from the orbifold locus of the critical manifold, and (H-MB) covers the rest (v1 claimed full verification; Proposition 5.3 shows why that was wrong).
(5) (v2) The quantitative compatibility (H-CONST) of Balaban's κ with the required window in γ is open, and the parameter exhibit of the suite shows it fails for typical $O(1)$ values of κ.
(6) (v2) The Hamiltonian form at the terminal scale is assumed ((H-HAM)), not derived.
(7) (v2) Theorem 3.1 is a package idealization of [1, 2, 3, 4, 5, 6] ((H-BAL)).

9.4. **Relation to the companion programme papers (new in v2).** This paper is the weak-coupling lattice attempt of the programme. The companion paper [17] (gradient flow) develops an all-β spectral framework whose central hypotheses (complete monotonicity (H0), flow monotonicity B′) are logically independent of the present route; [18] (framework) organizes the anomaly-algebra and quantum-information constraints. The three share the discipline of stating imported inputs as explicit hypotheses. The present paper's (H-CONST) plays a role analogous to [17]'s Hypothesis 1.1: a sharply isolated quantitative statement on which the respective conditional theorems pivot. Verification suites for all three live in the same repository (Appendix A).

Note added in v2

This version was prepared together with a numerical verification suite (`verify_2602_0032.py`, repository below). Changes relative to v1, all documented in situ: (1) title and abstract retagged from "self-contained proof" to "conditional reduction", with the four standing hypotheses of Section 1.3; (2) Proposition 5.3 (new): Morse–Bott failure at the orbifold locus of $\mathcal{M}_{\mathrm{flat}}$, exhibited by a non-perturbative finite-difference Hessian on a real lattice (kernel dimension 30 vs. $\dim T\widetilde{\mathcal{M}}_{\mathrm{flat}} = 24$ at $\theta = 0$, quartic ratio $15.99 \approx 2^4$, restoration to 26 at generic θ); consequently

(H2)/(H4) of Section 6 are amended and (H-MB) introduced; (3) Lemma 5.2: the v1 definition of $V_{\rm BO}$ (literal (2) inside (3)) is numerically non-positive and is replaced by the dispersion the v1 proof actually used; the Hessian constant is corrected to $N_c S_1$ (factor 2), improving $\omega_{\min}$ by $\sqrt{2}$; (4) Theorem 7.2 restated with the finite-window extraction and explicit error, correcting v1's inverted regime and its $M \to \infty$ limit (Remark 7.4); (5) Remark 7.3 revised: the (κ, γ) window is an open quantitative hypothesis, not a consequence of Balaban's construction; (6) Theorem 3.1(c): coupling-flow sign and normalization corrected (no downstream change); (7) Remark 4.3: sharp Ricci constant $N_c/2$; Section 4 flagged as unused context; (8) cross-references to the companion papers added. All v1 numbered statements and equations (1)–(7) keep their numbers. None of the above constitutes an unconditional proof, and the Clay problem [14] concerns the continuum theory in infinite volume, which remains out of scope.

Appendix A. Numerical verification

The suite `verify_2602_0032.py` (Python 3, NumPy only; repository `THE-ERIKSSON-PROGRAMME`, directory `verification/2602-0032/`) runs in a few seconds and performs 18 checks, all passing:

Block	Content	Status
1	Root identity $\sum_a \alpha_a \otimes \alpha_a = 2N_c I$, SU$(2\ldots5)$	verified (exact)
1	Hess $V_{\rm BO} = N_c S_1(L) I$ for (3) (FD, 3 cases)	verified (10^{-6})
1	v1's literal $V_{\rm BO}$: Hessian < 0, potential < 0	*refuted* (F2)
2	Hess $S_W = (1/N_c)(d^T d) \otimes I_3$ entrywise (SU(2), 2^3)	verified
2	$\theta = 0$: ker $= 30 > 24$; quartic ratio 16; generic θ: ker $= 26$	failure exhibited (F1)
3	$\mathrm{Ric}_{\mathrm{SU}(N_c)} = (N_c/2)\|X\|^2$ $(N_c = 2, 3)$	corrected (favorable)
4	doubling: windowed extraction within bound; $M > \kappa/m$ collapse	corrected (F3)
5	(κ, γ) window empty for $\kappa \le 5$, $h_0 \le 0.1$, $N_c = 2, 3$	exhibited (F4)

Run: `python3 verify_2602_0032.py` (exit code 0 iff all checks pass).

References

[1] T. Balaban, *Propagators and renormalization transformations for lattice gauge theories. I*, Comm. Math. Phys. **95** (1984), 17–40.
[2] T. Balaban, *Averaging operations for lattice gauge theories*, Comm. Math. Phys. **98** (1985), 17–51.
[3] T. Balaban, *Renormalization group approach to lattice gauge field theories. I*, Comm. Math. Phys. **109** (1987), 249–301.
[4] T. Balaban, *Renormalization group approach to lattice gauge field theories. II*, Comm. Math. Phys. **116** (1988), 1–22.
[5] T. Balaban, *Large field renormalization. I*, Comm. Math. Phys. **122** (1989), 175–202.
[6] T. Balaban, *Large field renormalization. II*, Comm. Math. Phys. **122** (1989), 355–392.
[7] B. O'Neill, *The fundamental equations of a submersion*, Michigan Math. J. **13** (1966), 459–469.
[8] A. L. Besse, *Einstein Manifolds*, Springer, 1987.
[9] L. Bérard-Bergery and J.-P. Bourguignon, *Laplacians and Riemannian submersions with totally geodesic fibres*, Illinois J. Math. **26** (1982), 181–200.
[10] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. **169** (2009), 903–991.
[11] B. Helffer and J. Sjöstrand, *Puits multiples en mécanique semi-classique. IV*, Comm. PDE **10** (1985), 245–340.
[12] D. Le Peutrec, F. Nier, and C. Viterbo, *Precise Arrhenius law for p-forms*, Ann. Henri Poincaré **14** (2013), 543–600.
[13] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, *QCD and instantons at finite temperature*, Rev. Mod. Phys. **53** (1981), 43–80.
[14] A. Jaffe and E. Witten, *Quantum Yang–Mills theory*, Clay Mathematics Institute, 2000.

[15] M. Lüscher, *Some analytic results concerning the mass spectrum of Yang–Mills gauge theories on a torus*, Nucl. Phys. B **219** (1983), 233–261.

[16] J. Koller and P. van Baal, *A non-perturbative analysis in finite volume gauge theory*, Nucl. Phys. B **302** (1988), 1–64.

[17] L. Eriksson, *Gradient flow monotonicity and the Yang–Mills mass gap*, ai.viXra:2602.0020 (v2), 2026.

[18] L. Eriksson, *Yang–Mills existence and mass gap: a framework via anomaly algebra, gradient-flow spectral methods, and quantum information*, ai.viXra:2602.0021 (v2), 2026.

Independent researcher
Email address: lluiseriksson@gmail.com