A machine-checked two-colour polymer gas on $\mathbb{Z}^4$: encoding, adjacency degrees, the box stage, and the Kotecký–Preiss condition down to Dobrushin

Kiichi Shiori Rin

2026-10-30 (revised 2026-11-14; text revised 2026-11-15)

wvbks0 [at] gmail.com · `https://computoergosum.com/en/index.html`

AI use is disclosed in §13; what is not claimed is in §11.

Abstract

We describe a Lean 4 formalisation of the combinatorial half of a cluster-expansion argument for a gas of connected clusters of plaquettes on $\mathbb{Z}^4$. A cluster X is encoded as a two-colour set: its *thin* plaquettes and its *defect* links. A defect link is a link of which X contains at least m of the six plaquettes, and each defect is labelled by $X \cap D_\ell$. Lean proves that the encoding preserves connectedness, that a cluster is recovered from its code and labels, and that the label sum is bounded by a product over links. The lattice objects are represented by concrete types: points $\texttt{Fin}\,4 \to \mathbb{Z}$, plaquettes = point $\times\,\texttt{Fin}\,6$, links = point $\times\,\texttt{Fin}\,4$. On these types the kernel decides the two-colour degrees 20, 44, 66, 122 and the 20 box shifts between link-sharing plaquettes (boxes of 16 points and at most 96 plaquettes). The Kotecký–Preiss per-point sum is then bounded on the plaquette and box stages and carried down to Dobrushin's condition and to $Z > 0$, $\log Z \geq -\sum a$. Activities enter only through hypotheses: a product bound, or the activity-sum hypothesis (H-sum) $\sum_{X \ni x} |z(X)| e^{2|X|} \leq 1$. No value or estimate of any activity is given. An identity rewrites the per-box sum of the box gas as a sum over plaquette clusters, so that (H-sum) in the box form yields Dobrushin's condition for the box gas in one composed theorem. The bounds on Z stated there hold for the box gas without (H-sum), its activities being non-negative. A bridge bounds the activity of the clusters meeting a box by plaquette-stage sums. Thirteen files contain 518 theorem declarations; after copies are removed, 163 distinct statements remain. Of these, 16 are finite statements closed by `decide +kernel` and the rest are proofs.

1 Introduction

Cluster expansions for lattice gauge theories at weak coupling separate the configuration into a region where every plaquette variable is small and a set of *large-field* plaquettes. They then treat the connected components of the large-field set as polymers of a gas whose activities are small (for the setting of the Yang–Mills problem see [2]). The analytic work lies in bounding the activities. A second, purely combinatorial, half is needed as well. One has to count connected clusters on the lattice with enough precision that a smallness condition on a single plaquette (or a single block) turns into the Kotecký–Preiss condition [3] for the whole gas, and this condition in turn into Dobrushin's condition [1] and lower bounds on the partition function. This note is about a version of that combinatorial half, written in Lean 4 against Mathlib [4], for one concrete organisation of the clusters that arose in the authors' work on a lattice $SU(2)$ model in four dimensions. The gas treated here is not the gas of connected components of a large-field set. Its polymers are connected clusters, and two of them are

compatible when their supports are disjoint, whereas distinct components of one set are disjoint and also non-adjacent. The per-point Kotecký–Preiss sums (§§3–5 and the first half of Theorem 5.1) are counts that do not depend on the compatibility relation. The passage to (D) and to the partition function (the second half of Theorem 5.1, and §7) concerns the gas as defined here.

The organisation is *two-colour*. Plaquettes that sit around a common link are strongly correlated, so a cluster that contains many of the six plaquettes of one link is better counted as a single object, a *defect link*, than as many independent plaquettes. A cluster is therefore encoded as a set of thin plaquettes together with a set of defect links, each defect link labelled by the subset of its six plaquettes that the cluster contains. The KP sum is then a sum over connected sets of a two-colour graph with colour-dependent weights and degrees. The note records what is formalised (§§2–7), how the lattice objects are represented (§5), which statements are decided by the kernel and which are proved (§8), and which inputs remain assumptions (§9). The authors' working records (unpublished) contain the numerical exploration that motivated the choices. None of it is used here.

Status of the results. Every implication stated as a theorem below is a Lean 4 theorem that passes its check. There is no `sorry`, no `axiom` declaration and no `native_decide`; the only axioms used are `propext`, `Classical.choice` and `Quot.sound`. The finite counts on $\mathbb{Z}^4$ are decided by `decide +kernel`. The activities of a gauge model are *not* treated. No upper bound on the activities is proved for a model. Every such bound is a hypothesis or is derived from one; the hypotheses are in product form (H-2col), in summed form (H-sum) (§2), or conditions named in the theorem that uses them. That the large-field clusters of a gauge model form a polymer gas of the kind considered here, with compatibility given by disjoint supports, is an assumption stated in §9 and is not formalised. The gas considered here is not the gas of connected components of a large-field set (§1). Where we say "proved" we mean "proved in Lean"; the rest of the text is explanation.

2 The setting: clusters, activities, and the two conditions

Graphs and connected sets. Let P be a type with decidable equality and adj a relation on P (symmetry is never used). For a finite $X \subseteq P$, Conn_X is the reflexive–transitive closure of $\{(a,b) : a,b \in X,\ \mathrm{adj}(a,b)\}$. The set X is *connected* if $\mathrm{Conn}_X(a,b)$ for all $a,b \in X$. A finite C is *rooted* at R if $R \subseteq C$ and every $c \in C$ is reached from some $\rho \in R$ inside C. For weights $x : P \to [0,\infty)$ put

$$W_g(U,R) = \sum_{C \subseteq U,\ C \text{ rooted at } R} \ \prod_{c \in C} x(c)$$

(`Wg`).

The gas. A *cluster* is a finite connected set X of plaquettes (adjacent = distinct and sharing a link) inside a finite region Λ_F. Let $\mathcal{X}$ be a finite family of clusters and $z : \mathcal{X} \to \mathbb{R}$ the activity. Two clusters are incompatible when their supports meet. Let supp be the support map: the identity on the plaquette stage, the set of boxes on the box stage (§6). With $\mu(X) = |z(X)| e^{2|\operatorname{supp} X|}$ we use:

- the *Kotecký–Preiss per-point condition* (a per-point form sufficient for the condition of [3], not that condition itself; the hypothesis `hsite` of `KPToDobrushin`):

$$\forall x : \quad \sum_{X \in \mathcal{X},\ x \in \operatorname{supp} X} |z(X)|\, e^{2|\operatorname{supp} X|} \ \le\ 1; \tag{KP}$$

- *Dobrushin's condition* with $a(X) = \log(1 + \mu(X))$:

$$|z(X)| \;\le\; (1 - e^{-a(X)}) \exp\Big(- \sum_{X' \neq X,\ \operatorname{supp} X' \cap \operatorname{supp} X \neq \emptyset} a(X')\Big). \tag{D}$$

The two activity hypotheses. Fix a set Λ_L of links and a threshold $m \ge 1$. For a link ℓ let D_ℓ be the set of its six plaquettes. For a cluster X,

$$\operatorname{def}(X) = \{\ell \in \Lambda_L :\ |X \cap D_\ell| \ge m\}, \qquad \operatorname{thin}(X) = \{p \in X :\ p \notin D_\ell \text{ for all } \ell \in \operatorname{def}(X)\}.$$

- **(H-2col)** (product form; the hypothesis `hz`): $|z(X)| \le \varepsilon_0^{|\operatorname{thin} X|} \prod_{\ell \in \operatorname{def} X} \eta(\ell, X \cap D_\ell)$ for numbers $\varepsilon_0 \ge 0$ and $\eta \ge 0$, together with a bound on the *label sum* of each link (the hypothesis `hlink`), $\sum_{S \subseteq D_\ell} \eta(\ell, S) e^{2|S|} \le c_L$ on the plaquette stage, $\sum_{S \subseteq D_\ell} \eta(\ell, S) \le \varepsilon_L$ on the box stage.
- **(H-sum)** (summed form): the inequality (KP) itself, assumed for the gas, either with $\operatorname{supp} = \mathrm{id}$ (plaquette stage) or with supp the box support of §6 (box stage).

(H-2col) is the form needed to *derive* (KP) from per-object smallness. (H-sum) is the form in which (KP) is taken as the assumption and only the passage to the box gas and to (D) is formalised. Where a theorem quoted below contains numerical constants in its hypotheses (for instance $\varepsilon_0 \le 2.5 \cdot 10^{-3}$), these are the thresholds of a sufficient condition proved in Lean. These constants are not estimates of any activity.

3 Two-colour connected sets and the per-point sum

The counting rests on a weighted form of the classical tree-graph/generating-function bound for connected sets.

Lemma 3.1 (Lean: `Wg_le`)**.** *Let* $\mathrm{nb}(p)$ *be finite sets with* $p \neq q \wedge \mathrm{adj}(p,q) \Rightarrow q \in \mathrm{nb}(p)$*, and let* $x, \tau \ge 0$ *satisfy* $x(p) \prod_{q \in \mathrm{nb}(p)} (1 + \tau(q)) \le \tau(p)$ *for all* p*. Then* $W_g(U, R) \le \prod_{r \in R} \tau(r)$ *for all finite* U, R*.*

The proof is a strong induction on $|U|$: removing one root r replaces it by a subset Q of its neighbours outside R (`rooted_erase`, `Wg_step`). The hypothesis on nb excludes $q = p$. This weakening, relative to the first version of the file, lets a defect link be adjacent to itself in the two-colour graph without entering its own neighbourhood. For a single colour and constant weights the same induction gives the count $\#\{C \ni p \text{ connected}, |C| = n\} \le e^n d^{n-1}$ when every point has at most $d \ge 1$ neighbours (`connected_count_le`), and the one-colour KP sum $\sum_{C \ni p} |z(C)| e^{2|C|} \le 1/d$ under $e^3 d\varepsilon \le 1$ (`point_kp_sum_le`).

Now colour the points, $\mathrm{col} : P \to \{0, 1\}$, and let d_{ab} bound the number of colour-b neighbours of a colour-a point. In Lean a colour is a `Bool`: `false` = thin plaquette, `true` = defect link.

Proposition 3.2 (Lean)**.** *Let* $\varepsilon_b, \theta_b \ge 0$ *and* $a_b \in \mathbb{R}$*. If* $|z(C)| \le \prod_{q \in C} \varepsilon_{\mathrm{col}(q)}$ *for every finite* C *and*

$$\varepsilon_b \, e^{a_b} (1 + \theta_0)^{d_{b0}} (1 + \theta_1)^{d_{b1}} \le \theta_b \qquad (b = 0, 1),$$

then for every finite U*, every point* p *and every finite root set* S*, with* C *ranging over the connected subsets of* U*,*

$$\sum_{C \ni p \ \mathit{conn.}} |z(C)| \prod_{q \in C} e^{a_{\mathrm{col}(q)}} \le \theta_{\mathrm{col}(p)}, \qquad \sum_{C \cap S \neq \emptyset \ \mathit{conn.}} |z(C)| \prod_{q \in C} e^{a_{\mathrm{col}(q)}} \le \sum_{r \in S} \theta_{\mathrm{col}(r)}.$$

(`two_color_point_sum`, `two_color_root_set_sum`; the version with point weights $x \le \xi_{\mathrm{col}}$ *in place of* z *is `two_color_root_gf`.)*

Lemma 3.3 (Lean). *With* $(d_{00}, d_{01}, d_{10}, d_{11}) = (20, 48, 72, 122)$, $(a_0, a_1) = (2, 12)$ *and* $(\theta_0, \theta_1) = (1/20, 10^{-4})$, *the two inequalities of Proposition 3.2 hold whenever* $\varepsilon_0 \le 2.5 \cdot 10^{-3}$ *and* $\varepsilon_1 \le 1.5 \cdot 10^{-11}$; *moreover* $\theta_0 + 4\theta_1 \le 1$ *(`lattice_fix`, `lattice_face_total`).*

The proof uses $e \le 2.7182818286$ (`exp_one_le`) and `norm_num`. The degrees 48 and 72 are upper bounds; the exact degrees 44 and 66 are decided in §5 and enter through "$\le$". The weight $e^{a_1} = e^{12} = (e^2)^6$ is the KP weight of the at most six plaquettes of a defect.

4 The encoding

Let F (plaquettes) and L (links) be types with decidable equality, $D : L \to \mathrm{Finset}\, F$, adj_F a relation on F, Λ_L a finite set of links and $m \ge 1$. The code of X is

$$\mathrm{enc}(X) = \mathrm{thin}(X) \sqcup \mathrm{def}(X) \;\subseteq\; F \oplus L,$$

and the two-colour adjacency `cadj` on $F \oplus L$ is

$$\begin{aligned}
&p \sim q \iff \mathrm{adj}_F(p,q),\\
&p \sim \ell \iff p \notin D_\ell \;\wedge\; \exists q \in D_\ell,\ \mathrm{adj}_F(p,q), \qquad \ell \sim p \iff p \notin D_\ell \;\wedge\; \exists q \in D_\ell,\ \mathrm{adj}_F(q,p),\\
&\ell \sim \ell' \iff \exists p \in D_\ell,\ \exists q \in D_{\ell'},\ p = q \;\vee\; \mathrm{adj}_F(p,q).
\end{aligned}$$

Lemma 4.1 (Lean, `LatticeTwoColorKP`). *For every finite* $X \subseteq F$:

*(i) (*connectedness, *`enc_connected`) if* X *is connected for* adj_F *and* $m \ge 1$, *then* $\mathrm{enc}(X)$ *is connected for `cadj`;*

*(ii) (*recovery, *`decode`, `enc_eq`, `thin_disjoint`)* $X = \mathrm{thin}(X) \cup \big(X \cap \bigcup_{\ell \in \mathrm{def}\, X} D_\ell\big)$; *equal codes have equal thin parts and equal defect sets; and* $\mathrm{thin}(X)$ *is disjoint from* $\bigcup_{\ell \in \mathrm{def}\, X} D_\ell$;

*(iii) (*size, *`card_le_enc`)* $|X| \le |\,\mathrm{thin}\, X| + \sum_{\ell \in \mathrm{def}\, X} |X \cap D_\ell|$;

*(iv) (*label sum, *`label_sum_le`) for* $g \ge 0$ *and any finite* $R \subseteq L$,

$$\sum_{T \subseteq \bigcup_{\ell \in R} D_\ell} \prod_{\ell \in R} g(\ell, T \cap D_\ell) \;\le\; \prod_{\ell \in R} \sum_{S \subseteq D_\ell} g(\ell, S),$$

even when the D_ℓ *overlap.*

The connectedness proof maps each plaquette $f \in X$ to its image $\mathrm{img}(f)$: a defect link containing it (chosen with `Classical.choose`) or f itself (`img_spec`). Adjacent plaquettes have adjacent or equal images (`img_adj`), so paths in X map to paths in $\mathrm{enc}(X)$. Every point of $\mathrm{enc}(X)$ is the image of, or adjacent to the image of, a plaquette of X. Injectivity is used in the following form. On the fibre $\{X : \mathrm{enc}(X) = C\}$ all clusters share $\mathrm{thin}(X)$ and $\mathrm{def}(X)$, so by (ii) the map $X \mapsto X \cap \bigcup_{\ell \in \mathrm{def}} D_\ell$ is injective there. This is the step `hinj` inside the proof of the next proposition, combined with (iv).

Proposition 4.2 (Lean: `encoded_sum_le`). *Under (H-2col), for any predicate* Q *satisfied by all codes* $\mathrm{enc}(X)$, $X \in \mathcal{X}$,

$$\sum_{X \in \mathcal{X}} |z(X)| e^{2|X|} \;\le \sum_{\substack{C \subseteq \Lambda_F \sqcup \Lambda_L \\ Q(C),\ C \text{ connected}}} \prod_{q \in C} x_w(q),$$

with connectedness for `cadj`, $x_w(p) = \varepsilon_0 e^2$ *and* $x_w(\ell) = \sum_{S \subseteq D_\ell} \eta(\ell, S) e^{2|S|}$.

The version with weight $e^{c|X|}$ for any $c \ge 0$ is `encoded_sum_le_c`; it is used with $c = 0$ on the box stage.

5 The lattice $\mathbb{Z}^4$ and the adjacency degrees

Representation. Points, plaquettes and links are

```
abbrev Pt := Fin 4 → ℤ
abbrev Face := Pt × Fin 6
abbrev Link := Pt × Fin 4
```

The plaquette index $k = 0, \dots, 5$ stands for the plane $(0,1), (0,2), (0,3), (1,2), (1,3), (2,3)$ and the first component is the base point. There are two tables at the origin. The table `links0` gives the four links of the plaquette $(0,k)$; for the plane (μ,ν) these are $(0,\mu), (e_\mu,\nu), (e_\nu,\mu), (0,\nu)$. The table `faces0` gives the six plaquettes containing the link $(0,\mu)$, namely $(0,\{\mu,\nu\})$ and $(-e_\nu,\{\mu,\nu\})$ for $\nu \neq \mu$. The embeddings `tF`, `tL` translate these tables to give `linksOf` and `facesOf` $= D$. Their consistency, $f \in D_\ell \iff \ell \in \mathrm{linksOf}(f)$, is `mem_facesOf_iff`; it is reduced by translation to the origin and proved there by case analysis (`key0`). Plaquette adjacency `faceAdj` means: distinct, and both in some D_ℓ.

Neighbourhoods and degrees. The two-colour neighbourhoods are translates of finite sets at the origin (writing ∂f for the links of f and $N_{FF}(f)$ for the translate of N^0_{FF}):

$$
\begin{aligned}
N^0_{FF}(k) &= \Big(\bigcup_{\ell \in \partial(0,k)} D_\ell \Big) \setminus \{(0,k)\},\\
N^0_{FL}(k) &= \Big\{ \ell \in \bigcup_{g \in N^0_{FF}(k)} \partial g : \ (0,k) \notin D_\ell \Big\},\\
N^0_{LF}(\mu) &= \Big(\bigcup_{f \in D_{(0,\mu)}} N_{FF}(f) \Big) \setminus D_{(0,\mu)},\\
N^0_{LL}(\mu) &= \Big(\bigcup_{f \in E_\mu} \partial f \Big) \setminus \{(0,\mu)\},
\end{aligned}
$$

with $E_\mu = D_{(0,\mu)} \cup \bigcup_{f \in D_{(0,\mu)}} N_{FF}(f)$. Lean shows that every `cadj`-neighbour (other than the point itself) lies in the translated set (`adj_mem_nbFF`, `fl_mem`, `lf_mem`, `ll_mem`, assembled in `lattice_hnb`) and decides the sizes:

```
theorem card_nbFF0 : ∀ k, (nbFF0 k).card = 20 := by decide +kernel
theorem card_nbFL0 : ∀ k, (nbFL0 k).card = 44 := by decide +kernel
theorem card_nbLF0 : ∀ μ, (nbLF0 μ).card = 66 := by decide +kernel
theorem card_nbLL0 : ∀ μ, (nbLL0 μ).card = 122 := by decide +kernel
theorem card_links0 : ∀ k, (links0 k).card = 4 := by decide +kernel
```

So a plaquette has 20 plaquette neighbours and 44 defect-link neighbours, and a defect link has 66 plaquette neighbours and 122 defect-link neighbours. Every plaquette has 4 links (`lattice_hfour`: at most 4 links of Λ_L contain a given plaquette). These close the hypotheses `hnb`, `hdeg` (`lattice_hdeg`, against the upper bounds $(20, 48, 72, 122)$ of Lemma 3.3) and `hfour` of the abstract theorem, leaving only the activity hypotheses:

Theorem 5.1 (Lean: `lattice_kp_site`, `lattice_dobrushin`)**.** *On $\mathbb{Z}^4$, under (H-2col) with $\varepsilon_0 \le 2.5 \cdot 10^{-3}$ and label sums $\le 1.5 \cdot 10^{-11} e^{12}$, (KP) holds on the plaquette stage; if moreover $|X| \ge 1$ and $|z(X)| e^{2|X|} \le e - 1$ for all $X \in \mathcal{X}$, then (D) holds for every $X \in \mathcal{X}$.*

```
theorem lattice_kp_site (ΛF : Finset Face) (ΛL : Finset Link) (m : ℕ) (hm : 1 ≤ m)
    (𝒳 : Finset (Finset Face)) (h𝒳 : ∀ X ∈ 𝒳, X ⊆ ΛF ∧ IsConnectedSet faceAdj X)
    (z : Finset Face → ℝ) (ε₀ : ℝ) (hε₀ : 0 ≤ ε₀) (hε₀' : ε₀ ≤ 2.5e-3)
    (η : Link → Finset Face → ℝ) (hη : ∀ ℓ S, 0 ≤ η ℓ S)
    (hlink : ∀ ℓ, ∑ S ∈ (facesOf ℓ).powerset, η ℓ S * Real.exp (2 * S.card) ≤ 1.5e-11
      * Real.exp 12)
    (hz : ∀ X ∈ 𝒳, |z X| ≤ (∏ _p ∈ thin facesOf ΛL m X, ε₀) *
      ∏ ℓ ∈ defects facesOf ΛL m X, η ℓ (X ∩ facesOf ℓ)) :
    ∀ x : Face, ∑ X ∈ 𝒳.filter (fun s => x ∈ id s), |z X| * Real.exp (2 * (id
      X).card) ≤ 1 :=
```

The proof is `encoded_kp_site`: Proposition 4.2 with $Q(C) =$ "C meets the root set of x", where the root set of a plaquette x is $\{x\} \sqcup \{\ell \in \Lambda_L : x \in D_\ell\}$ (`rootSet`, `enc_meets_root`); Proposition 3.2 with Lemma 3.3 bounds the result by $\theta_0 + 4\theta_1 \le 1$.

A wider adjacency. The file `Lattice24KP` repeats the construction with plaquette adjacency enlarged by the four parallel plaquettes across an elementary cube (`cubeOpp0`, `faceAdj24`). The decided degrees are 24, 44, 66, 122 (`card_nbFF24_0`, ..., `card_nbLL24_0`). The arithmetic condition holds for $\varepsilon_0 \le 2.1 \cdot 10^{-3}$, $\varepsilon_1 \le 1.5 \cdot 10^{-11}$ with $\theta = (1/23, 10^{-4})$ (`lattice24_fix`), and the analogue of Theorem 5.1 is `lattice_kp_site24`. At $\varepsilon_0 \ge 2.5 \cdot 10^{-3}$ no $\theta \ge 0$ satisfies the thin inequality with degrees $(24, 48)$ (`face24_fails`).

6 The box stage

Boxes and shifts. Let $\mathrm{blk}(x) = (\lfloor x_i/2 \rfloor)_i$ (`blk`, integer division). The box of $b \in \mathbb{Z}^4$ is $\{x : \mathrm{blk}(x) = b\} = 2b + \{0,1\}^4$. This box has 16 points (`card_corners`), 96 plaquettes based in it (`boxFaces`, `card_boxFaces`: ≤ 96) and 64 links (`card_boxLinks`: ≤ 64). A plaquette belongs to the box of its base point. The set of box shifts between plaquettes that share a link is computed at the origin, over the 16 positions of a base point in its box, and decided:

```
theorem card_boxOff : (boxOff.erase 0).card = 20 := by decide +kernel
```

`box_offset` shows that every link-sharing pair of plaquettes has its box shift in `boxOff`. Hence the box adjacency has degree ≤ 20 (`box_degree`) and the box image of a connected cluster is connected (`box_image_connected`). The 20 shifts are not the 8 nearest-neighbour shifts alone: `box_diag` exhibits two plaquettes sharing the link $((2,0,0,0),2)$ whose boxes differ by $(1,-1,0,0)$. Enumerating `boxOff` outside Lean gives the 8 shifts $\pm e_\mu$ and the 12 shifts $e_\mu - e_\nu$; only the total 20 is decided in Lean.

The two-colour box gas. Points of the box gas are pairs (box, colour). Thin–thin shifts are the 20 above (`card_offFF`). For pairs involving a defect, Lean decides that all point shifts in $N^0_{FL}, N^0_{LF}, N^0_{LL}$ have coordinates in $[-2,2]$ (`nbFL0_bdd`, `nbLF0_bdd`, `nbLL0_bdd`). Hence, for these pairs, the box shifts lie in $\{-1,0,1\}^4$ (`blk_sub_mem_cube`, `card_cube81`). Lean uses the degrees $(20, 81, 81, 80)$ (`card_offc`, `box_hdeg`). These are upper bounds; the exact values $39, 39, 56$ found by enumeration outside Lean are not used. The box image bimg of a `cadj`-connected set is connected for the box adjacency (`bimg_connected`). With the product hypothesis $|z_B(\Gamma)| \le \prod_{q \in \Gamma} \varepsilon_{\mathrm{col}(q)}$ over boxes, the arithmetic condition holds for $\varepsilon_{\mathrm{thin}} \le 2.55 \cdot 10^{-3}$, $\varepsilon_{\mathrm{def}} \le 10^{-9}$, $\theta = (1/20, 10^{-6})$ (`box_fix`), giving (KP) with $\mathrm{supp}\,\Gamma$ the set of boxes (`box_kp_site`, `box_hsite`) and, under the size and smallness hypotheses of Theorem 7.1, (D) (`box_dobrushin`). At $\varepsilon_{\mathrm{thin}} \ge 3.068 \cdot 10^{-3}$, `box_fails` shows that no $\theta \ge 0$ satisfies the thin inequality.

From the plaquette stage to the box gas. Let $\mathrm{boxImage}(X) = \mathrm{bimg}(\mathrm{enc}\, X)$ and define the box activity as the sum over the fibre,

$$z_B(\Gamma) = \sum_{X \in \mathcal{X},\ \mathrm{boxImage}(X) = \Gamma} |z(X)|$$

(`zBox`). The box support of $\mathrm{boxImage}(X)$ consists of the boxes of the base points of the thin plaquettes and of the defect links of X (`boxSupp`). This box support need not contain every box that contains a plaquette of X. Accordingly the sums over a box b below run over the clusters whose box support contains b, and the compatibility of the box gas is not claimed to agree with that of the plaquette gas. Under (H-2col) with $\sum_{S \subseteq D_\ell} \eta(\ell, S) \le \varepsilon_L$, the fibre sum is bounded by a product over boxes with $\varepsilon_{\mathrm{box}}(\mathrm{thin}) = (1+\varepsilon_0)^{96} - 1$ and $\varepsilon_{\mathrm{box}}(\mathrm{def}) = (1+\varepsilon_L)^{64} - 1$ (`fiber_box_le`, `cell_prod_le`, `box_hz_from_face`), and `face_to_box_dobrushin` carries this to (D) for the box gas when $\varepsilon_0 \le 2.6 \cdot 10^{-5}$, $\varepsilon_L \le 1.5 \cdot 10^{-11}$ (`epsBox_numbers`), under the size and smallness hypotheses of Theorem 7.1. This route needs the product form over boxes. We keep it as a formal statement but do not rely on it (§9).

The route used instead takes (H-sum) as the assumption. The box-stage sum of the box gas is *equal* to a sum over plaquette clusters:

Theorem 6.1 (Lean: `box_hsite_eq_face_sum`, `box_hsite_of_face_sum`)**.** *For every box b,*

$$\sum_{\Gamma \in \mathrm{boxImage}(\mathcal{X}),\ b \in \mathrm{supp}\,\Gamma} |z_B(\Gamma)|\, e^{2|\,\mathrm{supp}\,\Gamma|} = \sum_{X \in \mathcal{X},\ b \in \mathrm{supp}\,\mathrm{boxImage}(X)} |z(X)|\, e^{2|\,\mathrm{supp}\,\mathrm{boxImage}(X)|},$$

where $\mathrm{supp}\,\Gamma$ *is the set of boxes of* Γ*. Consequently (H-sum) in the box form, "the right-hand side is* ≤ 1 *for every b", is exactly (KP) for the box gas, with no product hypothesis.*

```
theorem box_hsite_eq_face_sum (ΛL : Finset Link) (m : ℕ) (𝒳 : Finset (Finset Face))
      (z : Finset Face → ℝ)
    (b : Pt) :
    ∑ Γ ∈ (𝒳.image (boxImage ΛL m)).filter (fun s => b ∈ boxSupp s),
        |zBox ΛL m 𝒳 z Γ| * Real.exp (2 * (boxSupp Γ).card) =
      ∑ X ∈ 𝒳.filter (fun X => b ∈ boxSupp (boxImage ΛL m X)),
        |z X| * Real.exp (2 * (boxSupp (boxImage ΛL m X)).card) := by
```

The proof uses only $|z_B| = z_B$ (a sum of absolute values) and regrouping by fibres (`sum_fiberwise_of_maps_to`). Feeding `box_hsite_of_face_sum` into `dobrushin_cond_of_kp` with supp = `boxSupp` gives (D) for the box gas, and Theorem 7.2 then gives the partition function. This composition is stated as one theorem, Theorem 7.3 in §7.

The bridge. The second use of the plaquette stage is to bound, box by box, the total activity of the clusters that touch a box. The root set of a box b is its 96 plaquettes together with the links of Λ_L having a plaquette in the box, at most $96 \cdot 4 = 384$ of them (`boxRoot`, `card_boxRoot_links`). The code of a cluster that meets the box meets this root set (`enc_meets_boxRoot`).

Theorem 6.2 (Lean: `box_event_bridge`)**.** *Under (H-2col) without KP weights, with label sums $\sum_{S \subseteq D_\ell} \eta(\ell, S) \le \varepsilon_L$, and if $\theta_0, \theta_1 \ge 0$ satisfy $\varepsilon_0(1+\theta_0)^{20}(1+\theta_1)^{48} \le \theta_0$ and $\varepsilon_L(1+\theta_0)^{72}(1+\theta_1)^{122} \le \theta_1$, then for every box b*

$$\sum_{X \in \mathcal{X},\ X \cap \mathrm{boxFaces}(b) \neq \emptyset} |z(X)| \ \le\ 96\,\theta_0 + 384\,\theta_1.$$

```
theorem box_event_bridge (ΛF : Finset Face) (ΛL : Finset Link) (m : ℕ) (hm : 1 ≤ m)
    (𝒳 : Finset (Finset Face)) (h𝒳 : ∀ X ∈ 𝒳, X ⊆ ΛF ∧ IsConnectedSet faceAdj X)
    (z : Finset Face → ℝ) (εo εL : ℝ) (hεo : 0 ≤ εo) (η : Link → Finset Face → ℝ) (hη
      : ∀ ℓ S, 0 ≤ η ℓ S)
    (hlink : ∀ ℓ, ∑ S ∈ (facesOf ℓ).powerset, η ℓ S ≤ εL)
    (hz : ∀ X ∈ 𝒳, |z X| ≤ (∏ _p ∈ thin facesOf ΛL m X, εo) *
      ∏ ℓ ∈ defects facesOf ΛL m X, η ℓ (X ∩ facesOf ℓ))
    (θ : Bool → ℝ) (hθ : ∀ c, 0 ≤ θ c)
    (hfix : ∀ a, (fun c => if c then εL else εo) a * (1 + θ false) ^ latticeDeg a
      false *
      (1 + θ true) ^ latticeDeg a true ≤ θ a) (b : Pt) :
    ∑ X ∈ 𝒳.filter (fun X => (X ∩ boxFaces b).Nonempty), |z X| ≤ 96 * θ false + 384 *
      θ true := by
```

The proof is `encoded_sum_le_c` with $c = 0$ and $Q(C) =$ "C meets the root set of b", followed by `two_color_root_gf` on $\mathbb{Z}^4$. In words: the total activity of the clusters that meet box b is controlled by the two-colour sums of the plaquette stage, without passing through a product over boxes. Turning this into a bound on the probability that a box contains a large plaquette requires a Peierls-type inequality $\mu(\text{box } b \text{ bad}) \le \sum_{X \cap b \neq \emptyset} \rho^*(X)$ for the measure; here μ denotes the measure, not the weight $\mu(X)$ of §2, and ρ^* is a cluster weight that this note does not define. This inequality is not formalised (§9). The file also contains two arithmetic instances of the hypotheses of Theorem 6.2 with literal inputs (`bridge_numbers_gamma`, `bridge_numbers_chernoff`). Their inputs are not discussed in this note.

7 From the Kotecký–Preiss sum to Dobrushin and to the partition function

Theorem 7.1 (Lean: `KPToDobrushin.dobrushin_cond_of_kp`)**.** *Let* $\operatorname{supp} : P \to \operatorname{Finset} S$*,* U *finite,* $|\operatorname{supp}\gamma| \ge 1$*, (KP) for* U*, and* $|z(\gamma)|e^{2|\operatorname{supp}\gamma|} \le e - 1$ *for* $\gamma \in U$*. Then (D) holds for every* $\gamma \in U$*.*

```
theorem dobrushin_cond_of_kp (supp : P → Finset S) (z : P → ℝ) (U : Finset P)
    (hsize : ∀ γ ∈ U, 1 ≤ (supp γ).card)
    (hsite : ∀ x, ∑ γ' ∈ U.filter (fun s => x ∈ supp s),
      |z γ'| * Real.exp (2 * (supp γ').card) ≤ 1)
    (hsmall : ∀ γ ∈ U, |z γ| * Real.exp (2 * (supp γ).card) ≤ Real.exp 1 - 1) :
    ∀ γ ∈ U, |z γ| ≤
      (1 - Real.exp (-Real.log (1 + |z γ| * Real.exp (2 * (supp γ).card)))) *
        Real.exp (-∑ γ' ∈ (U.erase γ).filter (fun s => (supp γ ∩ supp s).Nonempty),
          Real.log (1 + |z γ'| * Real.exp (2 * (supp γ').card))) := by
```

The proof bounds $a \le \mu$ (from $\log(1+t) \le t$). It then bounds the sum over polymers meeting $\operatorname{supp}\gamma$ by the sum over points of $\operatorname{supp}\gamma$ (`intersect_sum_le`), hence by $|\operatorname{supp}\gamma|$ by (KP). The remaining inequality $|z| \le (1 - e^{-a})e^{-n}$ is closed using $|z|e^{2n} \le e - 1$ and $n \ge 1$.

Theorem 7.2 (Lean: `PolymerGasDobrushin.dobrushin`, `Z_ge_exp`, `log_Z_ge`)**.** *For a finite gas* U *with symmetric incompatibility relation, real activities of any sign, and* $Z(\Gamma) = \sum_{t \subseteq \Gamma \text{ compatible}} \prod_{s \in t} z(s)$*: if* $a \ge 0$ *is any function and (D) holds with this* a *in place of* $\log(1+\mu)$ *and "incompatible" in place of "supports meet", then for every* $\Gamma \subseteq U$*,* $Z(\Gamma) > 0$*,* $e^{-a(\gamma)} Z(\Gamma \setminus \gamma) \le Z(\Gamma)$ *for* $\gamma \in \Gamma$*,* $Z(\Gamma) \ge \exp(-\sum_{\gamma \in \Gamma} a(\gamma))$ *and* $\log Z(\Gamma) \ge -\sum_{\gamma \in \Gamma} a(\gamma)$*.*

With the incompatibility relation "supports meet" the hypothesis of Theorem 7.2 is literally the conclusion of Theorem 7.1. For the box gas the three steps are composed in one file:

Theorem 7.3 (Lean: `BoxGasFromFaceSum.box_dobrushin_of_face_sum`, `box_gas_partition_of_face_sum`)**.** *Let $\mathcal{X}$ be a finite family of plaquette sets, z real, and suppose that every box image $\Gamma = \mathrm{boxImage}(X)$, $X \in \mathcal{X}$, has $|\operatorname{supp}\Gamma| \geq 1$, and that (H-sum) holds in the box form, $\sum_{X\in\mathcal{X},\ b\in\operatorname{supp}\mathrm{boxImage}(X)} |z(X)|e^{2|\operatorname{supp}\mathrm{boxImage}(X)|} \leq 1$ for every box b. Then the box gas (polymers $\mathrm{boxImage}(\mathcal{X})$, activity z_B, incompatibility "box supports meet") satisfies (D) with $a(\Gamma) = \log(1 + |z_B(\Gamma)|e^{2|\operatorname{supp}\Gamma|})$, and for every $\Gamma_s \subseteq \mathrm{boxImage}(\mathcal{X})$: $Z(\Gamma_s) > 0$, $Z(\Gamma_s) \geq \exp(-\sum_{\Gamma\in\Gamma_s} a(\Gamma))$ and $\log Z(\Gamma_s) \geq -\sum_{\Gamma\in\Gamma_s} a(\Gamma)$.*

```
theorem box_gas_partition_of_face_sum (ΛL : Finset Link) (m : ℕ) (𝒳 : Finset (Finset
      Face))
    (z : Finset Face → ℝ)
    (hsize : ∀ Γ ∈ 𝒳.image (boxImage ΛL m), 1 ≤ (boxSupp Γ).card)
    (hsum : ∀ b : Pt, ∑ X ∈ 𝒳.filter (fun X => b ∈ boxSupp (boxImage ΛL m X)),
      |z X| * Real.exp (2 * (boxSupp (boxImage ΛL m X)).card) ≤ 1) :
    ∀ Γs ⊆ 𝒳.image (boxImage ΛL m),
      0 < PolymerGasDobrushin.Z boxInc (zBox ΛL m 𝒳 z) Γs ∧
      Real.exp (-∑ Γ ∈ Γs, boxA ΛL m 𝒳 z Γ) ≤ PolymerGasDobrushin.Z boxInc (zBox ΛL m
      𝒳 z) Γs ∧
      -∑ Γ ∈ Γs, boxA ΛL m 𝒳 z Γ ≤ Real.log (PolymerGasDobrushin.Z boxInc (zBox ΛL m
      𝒳 z) Γs) := by
  have hcond := box_dobrushin_of_face_sum ΛL m 𝒳 z hsize hsum
```

Since $z_B \geq 0$ by definition, the bounds on Z in Theorem 7.3 hold for the box gas whether or not (H-sum) does: with non-negative activities $Z(\Gamma_s) \geq 1$ (the empty family contributes 1), while $\exp(-\sum a) \leq 1$ and $-\sum a \leq 0$. The content of Theorem 7.3 is therefore (D) for the box gas, obtained through the identity of Theorem 6.1. The bounds on Z are recorded because the composed file states them. The smallness hypothesis $|z_B(\Gamma)|e^{2|\operatorname{supp}\Gamma|} \leq e - 1$ of Theorem 7.1 is not assumed: by Theorem 6.1 the term of Γ in the box-stage sum at a box $b \in \operatorname{supp}\Gamma$ is at most $1 \leq e-1$. The file `BoxGasFromFaceSum.lean` consists of verbatim copies of `BoxFromFace.lean` and `PolymerGasDobrushin.lean`, each in its own namespace, followed by the two composed theorems and two auxiliary lemmas (symmetry of the incompatibility, $a \geq 0$).

Tools for checking (H-sum). `ClassSiteSum.hsite_of_class_sum` derives (KP) from $|z(\gamma)| \leq c\,\varepsilon_{\mathrm{cls}\gamma}^{|\operatorname{supp}\gamma|}$ and from counts $N(k,n)$ of clusters through a point by class and size, when $c\sum_k \sum_{n\leq M} N(k,n)\varepsilon_k^n e^{2n} \leq 1$. `SizeSplitKP.size_split_site_sum` bounds the clusters above a size n_0 by a geometric series when the counts are $\leq cK^{n-1}$ and $K\eta < 1$.

8 What is decided and what is proved

A statement is *decided* when its whole proof is `by decide +kernel`: the kernel evaluates the finite sets and checks the claim. These sets are built from `Finset` operations on `Fin`, $\mathbb{Z}$ and the tables `links0`, `faces0`. All other statements have tactic proofs. Table 1 counts theorem declarations per file and separates verbatim copies from new statements. It counts one declaration per `#print axioms` line. The two `@[simp]` lemmas `tF_apply`, `tL_apply`, proved by `rfl`, are not printed and not counted.

Counting each decided fact once, 16 distinct statements are decided: the face degrees $20, 44, 66, 122$ and the four links of a plaquette; the 20 box shifts (twice, as `card_boxOff` and `card_offFF`, built differently); $|\{-1,0,1\}^4| = 81$; the three coordinate bounds $|\cdot| \leq 2$; $|\{0,1\}^4| = 16$; and the degrees $24, 44, 66, 122$ of the wider adjacency. The remaining 147 distinct statements are proofs; one of them, `box_diag`, uses `decide` on concrete memberships

Table 1: Theorem declarations. "Copied": same name and statement as in the earlier file named, up to the weakening stated below; "decided": whole proof `by decide +kernel`.

File	theorems	copied	new	decided (new)
`ConnectedSetCounting`	14	–	14	–
`TwoColorPolymerSum`	18	3	15	–
`KPToDobrushin`	2	–	2	–
`TwoColorEncoding`	37	20	17	–
`LatticeTwoColorKP`	68	37	31	6
`BoxTwoColorKP`	71	51	20	5
`Lattice24KP`	73	53	20*	4
`BoxFromFace`	100	81	19	1
`PolymerGasDobrushin`	8	–	8	–
`ClassSiteSum`	4	–	4	–
`SizeSplitKP`	6	–	6	–
`BoxConnectivity`	5	–	5	–
`BoxGasFromFaceSum`	112	108	4	–
total	518	353	165	16

* Includes `quad_le_pow` and `exp_two_gt`, which also occur in `BoxTwoColorKP`; counted once among the 163 distinct statements. Copies: `TwoColorPolymerSum` §1 from `ConnectedSetCounting`; `TwoColorEncoding` §§1–6 from `TwoColorPolymerSum` and §10 from `KPToDobrushin`; `LatticeTwoColorKP` §§1–10 from `TwoColorEncoding` with the neighbourhood hypothesis weakened to $p \neq q$; `BoxTwoColorKP` and `Lattice24KP` §§1–11 from `LatticeTwoColorKP`; `BoxFromFace` §§1–12 from `BoxTwoColorKP`, and its §15 repeats the five decided degree counts of `LatticeTwoColorKP`; `BoxGasFromFaceSum` contains `BoxFromFace` and `PolymerGasDobrushin` whole.

inside a short proof. They include every inequality on sums, all statements about connectedness and the encoding, and all arithmetic with e (by `norm_num` after $e \leq 2.7182818286$ or $7.389 < e^2$). A decided count says that a set *defined in Lean* has the stated size. The membership lemmas (`adj_mem_nbFF`, `fl_mem`, `lf_mem`, `ll_mem`, `box_offset`), which are proofs, check that the definition matches the intended geometric neighbourhood in one direction only: every true neighbour lies in the set. Sets that are too large would only weaken the bounds.

9 What remains an assumption

The product form over boxes asks that the activity of a box configuration be bounded by a product of per-box factors; the authors' numerical exploration indicates that correlations between plaquettes selected in different boxes make it unsuitable for the model they have in mind. The identity of Theorem 6.1, with (H-sum) as the assumption, is therefore the route kept here. Whether (H-sum) holds for a particular gauge model is not addressed.

10 What is machine-checked

The thirteen files of Table 1 each begin with `import Mathlib` and are checked as single files. They are distributed as one bundle (ledger and build instructions: `https://computoergosum.com/en/principia/lean/index.html`; bundle: `https://computoergosum.com/principia/lean/principia-ym-two-colour-polymer-2026-10-30.tar.gz`, checksum file `https://computoergosum.com/principia/lean/principia-ym-two-colour-polymer-2026-10-30.tar.gz.sha256`; axiom log: `https://computoergosum.com/principia/lean/axioms-ym-two-colour-polymer-2026-10-30.txt`). To verify: extract the bundle; place the files in a Lake project pinned to Lean v4.33.1 and Mathlib v4.33.1 (commit `0df444a3`); run `lake exe cache get`; then

Table 2: Status of each ingredient.

Ingredient	Grade
encoding: connectedness, recovery, size, label and fibre sums	Lean
two-colour KP sum (Lemma 3.1, Proposition 3.2)	Lean
degrees $20, 44, 66, 122$; 4 links; 20 box shifts; 16 points	Lean (`decide +kernel`)
arithmetic sufficient conditions (Lemma 3.3, `box_fix`, `lattice24_fix`)	Lean (`norm_num`)
(KP) $\Rightarrow$ (D) $\Rightarrow$ $Z > 0$, $\log Z \geq -\sum a$	Lean
(H-sum) $\Rightarrow$ (D) for the box gas, Theorem 7.3 (its Z bounds hold without (H-sum))	Lean
identity of Theorem 6.1; bridge, Theorem 6.2	Lean
(H-2col) on the plaquette stage, label-sum bounds	assumption (hypotheses `hz`, `hlink`)
(H-sum)	assumption
product form over boxes (`box_hz_from_face`, `box_kp_site`)	assumption, not relied on
clusters form a gas, incompatibility = meeting supports (not the gas of connected components of a large-field set)	assumption, not formalised
Peierls-type bound $\mu(\text{box bad}) \leq \sum_{X \cap b \neq \emptyset} \rho^*(X)$	by hand, not formalised
box degrees $39, 39, 56$; split $8 + 12$ of the shifts	computed outside Lean, unused

run `LEAN_NUM_THREADS=1 lake env lean <file>.lean` for each file and compare the `#print axioms` lines at its end with `axioms-<file>.txt`. There is no `sorry`, no `axiom` declaration and no `native_decide`. The file carrying the main statements of §6, `BoxFromFace.lean` (namespace `BoxFromFace`, 100 theorems), has sha256 `fbab4754f910061a4581d78c80dd35f05335e3c82abc0a296e6d027f12016f69`; `LatticeTwoColorKP.lean` (Theorem 5.1) has `0f24fff2e02976f50574d54931cef829a1314fc54e2983b7e0931396dd6fa049`; `BoxGasFromFaceSum.lean` (Theorem 7.3) has `f659c624ac9f43d2a5a2a5e4e751cb6399a10c3496781dcaf2258f78fbb12b3c`. The checksums of all thirteen files are in the file `SHA256SUMS` of the bundle.

Every theorem depends on at most `propext`, `Classical.choice` and `Quot.sound`. The only exception in form is `blk_add` (in each file that contains it), which depends on `propext` and `Quot.sound` only. In the logs of `BoxFromFace`, `BoxGasFromFaceSum`, `Lattice24KP`, `LatticeTwoColorKP`, `BoxTwoColorKP`, `TwoColorEncoding`, `TwoColorPolymerSum` and `ConnectedSetCounting` the axiom is printed as `choice` rather than `Classical.choice`: these files contain `open Classical` (in `BoxGasFromFaceSum` twelve of the lines read `Classical.choice`), under which `#print axioms` shows the short name of the same constant. It is not a different axiom.

Lean checks the rows marked "Lean" in Table 2, in the generality of the hypotheses quoted. It does not check that `hz`, `hlink` or (H-sum) hold for any model, nor the polymer form of a gauge model's large-field expansion.

11 What is not claimed

- No bound on the activities of any lattice gauge model is claimed or derived. The numerical constants in the hypotheses ($2.5 \cdot 10^{-3}$, $1.5 \cdot 10^{-11}$, $2.55 \cdot 10^{-3}$, 10^{-9}, $2.6 \cdot 10^{-5}$, $2.1 \cdot 10^{-3}$) are thresholds of sufficient conditions, not estimates.

- (H-2col), the label-sum bounds and (H-sum) are assumptions. We do not claim that they hold for the $SU(2)$ model that motivated this work. The product form over boxes is kept only as a formal statement.

- That the large-field expansion of a gauge theory yields a polymer gas of the kind of §2 (compatibility by disjoint supports, activities attached to clusters) is not formalised.
- The degrees used for pairs involving a defect on the box stage are upper bounds $(81, 81, 80)$, not exact counts. The degrees 48, 72 in Lemma 3.3 are upper bounds for the decided 44, 66.
- The cluster-counting bounds are of standard type [3, 1, 5]. We do not know whether the particular two-colour encoding with labelled defect links has been noted before.

12 Correspondence table

Table 3: Statements of this note and Lean declarations. Where a declaration is copied into later files, the line of the file named is given.

Statement	Lean declaration	File:line
Lemma 3.1	`Wg_le`	LTK:148
count $\le e^n d^{n-1}$	`connected_count_le`	CSC:249
one-colour KP sum $\le 1/d$	`point_kp_sum_le`	CSC:420
Proposition 3.2, generating function	`two_color_Wg_le`	LTK:243
Proposition 3.2, point	`two_color_point_sum`	LTK:274
Proposition 3.2, root set	`two_color_root_set_sum`	LTK:311
Proposition 3.2, point weights	`two_color_root_gf`	LTK:833
Lemma 3.3	`lattice_fix`	LTK:397
Lemma 3.3, $\theta_0 + 4\theta_1 \le 1$	`lattice_face_total`	LTK:427
abstract face KP sum (first version)	`lattice_two_color_kp`	TPS:435
def, thin, enc, `cadj`	`defects`, `thin`, `enc`, `cadj`	LTK:539
Lemma 4.1(i)	`enc_connected`	LTK:582
Lemma 4.1(ii)	`decode`, `enc_eq`, `thin_disjoint`	LTK:629
Lemma 4.1(iii)	`card_le_enc`	LTK:663
Lemma 4.1(iv)	`label_sum_le`	LTK:490
Proposition 4.2	`encoded_sum_le`	LTK:715
Proposition 4.2, weight $e^{c\|X\|}$	`encoded_sum_le_c`	BFF:1523
abstract (KP), plaquette stage	`encoded_kp_site`	LTK:876
abstract (D), plaquette stage	`encoded_dobrushin`	LTK:1030
$f \in D_\ell \iff \ell \in \partial f$	`mem_facesOf_iff`	LTK:1107
plaquette adjacency	`faceAdj`	LTK:1111
membership in $N_{FF}, N_{FL}, N_{LF}, N_{LL}$	`adj_mem_nbFF`, `fl_mem`, `lf_mem`, `ll_mem`	LTK:1116
degrees $20, 44, 66, 122$; 4 links (decided)	`card_nbFF0`, `card_nbFL0`, `card_nbLF0`, `card_nbLL0`, `card_links0`	LTK:1171
`hnb` on $\mathbb{Z}^4$	`lattice_hnb`	LTK:1187
`hdeg` on $\mathbb{Z}^4$	`lattice_hdeg`	LTK:1211
`hfour` on $\mathbb{Z}^4$	`lattice_hfour`	LTK:1220
Theorem 5.1, (KP)	`lattice_kp_site`	LTK:1228
Theorem 5.1, (D)	`lattice_dobrushin`	LTK:1240
shift of adjacent plaquettes	`box_offset`	LTK:1280
20 box shifts (decided)	`card_boxOff`	LTK:1291
box degree ≤ 20	`box_degree`	LTK:1299
box image connected	`box_image_connected`	LTK:1310
diagonal shift $(1, -1, 0, 0)$	`box_diag`	LTK:1328
wider adjacency	`cubeOpp0`, `faceAdj24`	L24:1172
degrees $24, 44, 66, 122$ (decided)	`card_nbFF24_0` … `card_nbLL24_0`	L24:1237
wider adjacency, arithmetic	`lattice24_fix`	L24:1300
wider adjacency, (KP)	`lattice_kp_site24`	L24:1390
wider adjacency, sharpness	`face24_fails`	L24:1432
20 thin box shifts (decided)	`card_offFF`	BTK:1207
$\|\{-1, 0, 1\}^4\| = 81$ (decided)	`card_cube81`	BTK:1212

Statement	Lean declaration	File:line
coordinate bounds (decided)	`nbFL0_bdd`, `nbLF0_bdd`, `nbLL0_bdd`	BTK:1215
box degrees $(20, 81, 81, 80)$	`card_offc`	BTK:1243
box `hnb`, `hdeg`	`box_hnb`, `box_hdeg`	BTK:1258
box image of a two-colour set	`bimg_connected`	BTK:1310
box arithmetic	`box_fix`	BTK:1334
box (KP), root box	`box_kp_site`	BTK:1361
box sharpness	`box_fails`	BTK:1402
box (KP), supp = boxes	`box_hsite`	BTK:1432
box (D)	`box_dobrushin`	BTK:1464
16 points (decided)	`card_corners`	BFF:1662
fibre over a box set	`fiber_box_le`	BFF:1610
product form from (H-2col)	`box_hz_from_face`	BFF:1744
$\varepsilon_{\mathrm{box}}$ arithmetic	`epsBox_numbers`	BFF:1785
plaquette stage $\Rightarrow$ box (D)	`face_to_box_dobrushin`	BFF:1842
Theorem 6.1, identity	`box_hsite_eq_face_sum`	BFF:1807
Theorem 6.1, (H-sum) $\Rightarrow$ box (KP)	`box_hsite_of_face_sum`	BFF:1833
root set of a box, ≤ 384 links	`boxRoot`, `card_boxRoot_links`	BFF:1940
Theorem 6.2	`box_event_bridge`	BFF:1970
arithmetic instances (inputs not discussed)	`bridge_numbers_gamma`, `bridge_numbers_chernoff`	BFF:2009
sum over meeting polymers	`intersect_sum_le`	KPD:27
Theorem 7.1	`dobrushin_cond_of_kp`	KPD:46
Theorem 7.2, $Z > 0$	`dobrushin`	PGD:118
Theorem 7.2, lower bounds	`Z_ge_exp`, `log_Z_ge`	PGD:194
class sum $\Rightarrow$ (KP)	`hsite_of_class_sum`	CSS:74
size split	`size_split_site_sum`	SSK:80
Theorem 7.3, (D) for the box gas	`box_dobrushin_of_face_sum`	BGF:2396
Theorem 7.3, Z bounds	`box_gas_partition_of_face_sum`	BGF:2428
incompatibility, $a \geq 0$ (auxiliary)	`boxInc_symm`, `boxA_nonneg`	BGF:2384

File abbreviations: CSC `ConnectedSetCounting`, TPS `TwoColorPolymerSum`, TCE `TwoColorEncoding`, LTK `LatticeTwoColorKP`, BTK `BoxTwoColorKP`, L24 `Lattice24KP`, BFF `BoxFromFace`, KPD `KPToDobrushin`, PGD `PolymerGasDobrushin`, CSS `ClassSiteSum`, SSK `SizeSplitKP`, BCN `BoxConnectivity`, BGF `BoxGasFromFaceSum`. Namespaces equal file names. BCN (blocks of side L are boxes; a box with points in and out of A has an edge leaving A) is in the bundle but not used above.

13 Disclosure of AI use

Two of the three authors are AI agents. *Shiori* and *Rin* are instances of Claude (Anthropic), run as the coding agent Claude Code. The encoding and the Lean 4 files including the formalised statements were produced by Shiori over many sessions, which share no memory and rely on the written record; further sessions of Shiori reviewed the text against the Lean statements. Rin re-checked the Lean development from the distributed archive, confirmed that the declarations cited here are in it, and prepared the web edition. The text was first written by Rin from the files and reports and later rewritten by Shiori. Kiichi, the human author, set the task and the working rules and decides what is published; he is responsible for this note.

This note has not been refereed. All review was done by further instances of the same model family, which share the authors' blind spots. What the Lean kernel checks is checked; everything else is the authors' word. In particular, the hypotheses on the activities ((H-2col), the label-sum bounds and (H-sum)) are assumptions, the polymer form of a gauge model's large-field expansion is not formalised, and the Peierls-type bound is by hand (§9).

Revision history. Revised 2026-11-14: the relation of the gas treated here to the gas of connected components is stated; the bounds on Z in Theorem 7.3 are noted to be automatic for the box gas; the definition of the box support is spelled out; one count in a table footnote is corrected; a reference to background material outside the note is removed, so that the note stands alone. Revised 2026-11-15: the text was rewritten for readability. In addition, the abstract now says that (H-sum) is used in its box form and that the label sum is bounded by a product (not factorised); the size and smallness hypotheses of Theorem 7.1 are named where condition (D) for the box stage is quoted; Theorem 7.2 names the gas U; the paragraph on the status of the results says that every upper bound on the activities is a hypothesis or is derived from one, and in which forms the hypotheses come; and the count of decided statements, the caption of Table 1, the symbols of the Peierls-type inequality and the description of the axiom lines were made precise. The Lean statements, the proofs, the numbers and the bodies of the tables are unchanged.

References

[1] R. L. Dobrushin, *Estimates of semi-invariants for the Ising model at low temperatures*, in: Topics in Statistical and Theoretical Physics, Amer. Math. Soc. Transl. (2) 177 (1996) 59–81 (not consulted directly).

[2] A. Jaffe, E. Witten, *Quantum Yang–Mills theory*, in: The Millennium Prize Problems, Clay Mathematics Institute, 2006, 129–152 (not consulted directly).

[3] R. Kotecký, D. Preiss, *Cluster expansion for abstract polymer models*, Comm. Math. Phys. 103 (1986) 491–498 (not consulted directly).

[4] The mathlib Community, *The Lean mathematical library*, in: Proc. 9th ACM SIGPLAN Int. Conf. on Certified Programs and Proofs (CPP 2020), 367–381.

[5] D. Ueltschi, *Cluster expansions and correlation functions*, Moscow Math. J. 4 (2004) 511–522 (not consulted directly).