Graphons in Lean 4: Blueprint

Cameron Freer

March 21, 2026

Contents

Chapter 1

Graphons

This project formalizes the theory of graphons — limits of dense graph sequences — in Lean 4 using Mathlib, following Lovász [Lov12].

A *graphon* is a symmetric measurable function $W : [0,1]^2 \to [0,1]$ (more generally, on an arbitrary probability space). Graphons arise as the natural limit objects for convergent sequences of dense graphs.

1.1 Proof Status

Three remaining `sorry` declarations, driven by two main missing mathematical inputs (no custom axioms are introduced):

1. **Rokhlin's theorem** (`MeasurePreserving.exists_common_extension`): any two standard Borel probability spaces admit a common measure-preserving extension. Mathlib provides `PolishSpace.measurableEquiv` (the measurable isomorphism) but not the measure-preserving version; proving this is in progress. Used by: cut distance triangle inequality, partition alignment, compactness.

2. **Algebraic determination** for $k \geq 2$ (`matrix_quotient_of_weightedHomSum_eq`, positive-weight case): Lovász [Lov12, Theorem 5.30], quotient form. The $k = 1$ case is fully proved; the $k \geq 2$ case requires graph algebra separation arguments that are partially built. Used by: inverse counting lemma core.

3. **Determination pending theorem** (`cutDistance_zero_of_homDensity_eq`): depends on both of the above. Used by: convergence equivalence.

All other declarations contain no additional `sorry`s. These three `sorry`s will be replaced with proofs.

1.2 Main Results

The formalization establishes:

- **Cut distance pseudometric.** Cut distance is a pseudometric: non-negativity, symmetry (Proposition 7), and the triangle inequality (Theorem 8).

- **Frieze–Kannan weak regularity lemma.** Every graphon admits a step approximation of bounded complexity (Theorem 14), following Frieze–Kannan [FK99] and Lovász [Lov12, Corollary 9.13].
- **Counting lemma.** Small cut distance implies similar homomorphism densities (Theorem 16).
- **Compactness.** The cut-distance quotient modulo weak isomorphism is a compact metric space; concretely, the graphon pseudometric space is totally bounded (Theorem 17) and complete (Theorem 18).
- **Inverse counting lemma.** For any $\varepsilon > 0$, finitely many test graphs control cut distance up to ε (Theorem 19), following Lovász [Lov12, Lemma 10.32].
- **Convergence equivalence.** Cut distance convergence is equivalent to convergence of all homomorphism densities (Theorem 20).

1.3 Graphon Definition

A graphon on a probability space (α, μ) is a symmetric measurable function taking values in $[0, 1]$ a.e., following Lovász [Lov12, Definition 7.1].

Definition 1 (Graphon)**.**

Chapter 2

Cut Distance

The *cut distance* is the fundamental pseudometric on graphon space, measuring how close two graphons are after optimal rearrangement of the underlying probability space. The development follows Lovász [Lov12, Chapter 8] and Borgs–Chayes–Lovász–Sós–Vesztergombi [BCL+08].

2.1 Cut Norm

The cut norm of the difference of two graphons, defined as the supremum of $|\int_{S\times T}(U-W)|$ over measurable sets S, T.

Definition 2 (Cut norm difference)**.**

2.2 Pullback

The pullback of a graphon under a measure-preserving map, fundamental to the definition of cut distance. Following Lovász [Lov12, Section 8.2].

Definition 3 (Pullback of a graphon)**.**

2.3 Cut Distance Definition

The cut distance between two graphons, defined as the infimum of $\|U^\varphi - W^\psi\|_\square$ over measure-preserving maps φ, ψ. This two-sided formulation avoids the need for invertibility (Lovász [Lov12, Section 8.2]).

Definition 4 (Cut distance)**.**

2.4 Metric Properties

Cut distance is a pseudometric on graphon space (Lovász [Lov12, Theorem 8.13, Corollary 8.14]).

Theorem 5 (Non-negativity of cut distance)**.**

Proof. □

Theorem 6 (Cut distance to self is zero)**.**

Proof. □

Theorem 7 (Symmetry of cut distance)**.**

Proof. □

The triangle inequality uses Rokhlin's theorem (that standard Borel probability spaces are isomorphic) to align measure-preserving witnesses. *Status:* proved modulo a `sorry`'d Rokhlin interface (`exists_common_extension`), which is planned to be proved in future work.

Theorem 8 (Triangle inequality for cut distance)**.**

Proof. □

Pullback by a measure-preserving bijection preserves cut distance (Lovász [Lov12, Section 8.2]).

Theorem 9 (Pullback invariance of cut distance)**.**

Proof. □

Chapter 3

Partitions and Step Graphons

Step graphons — graphons that are constant on rectangles of a partition — are the finite-dimensional building blocks of graphon theory.

3.1 Measurable Partitions

A finite partition of the probability space into measurable sets that cover almost everywhere.

Definition 10 (Measurable partition)**.**

3.2 Stepification

The stepification W_P of a graphon W with respect to a partition P replaces W on each rectangle $S \times T$ by its average (Lovász [Lov12, Section 9.2]).

Definition 11 (Stepification of a graphon)**.**

3.3 Step Graphons from Coefficients

A step graphon built from explicit coefficients on a partition.

Definition 12 (Step graphon from coefficients)**.**

Chapter 4

Regularity

The Frieze–Kannan weak regularity lemma for graphons: every graphon can be approximated by a step graphon of bounded complexity. Following Frieze–Kannan [FK99] and Lovász [Lov12, Corollary 9.13].

4.1 Energy Increment

The energy increment lemma: if a partition P does not approximate W well, then a refinement Q has strictly higher energy (Lovász [Lov12, Lemma 9.11]).

Theorem 13 (Energy increment lemma)**.**

Proof. □

4.2 Regularity Lemma

The regularity lemma: for any $\varepsilon > 0$, every graphon admits a partition P with at most $\text{regularityBound}(\varepsilon)$ parts such that $\|W - W_P\|_\square \leq \varepsilon$.

Theorem 14 (Frieze–Kannan weak regularity lemma)**.**

Proof. □

Chapter 5

Homomorphism Densities and the Counting Lemma

Homomorphism densities encode the local structure of a graphon. The counting lemma shows that small cut distance implies similar homomorphism densities. Following Lovász [Lov12, Chapter 5 and Section 10.1].

5.1 Homomorphism Density

For a finite graph F and a graphon W, the homomorphism density $t(F, W) = \int \prod_{uv \in E(F)} W(x_u, x_v)\, d\mu^{V(F)}$ (Lovász [Lov12, Equation (5.29)]).

Definition 15 (Homomorphism density)**.**

5.2 Counting Lemma

The counting lemma: $|t(F, U) - t(F, W)| \leq |E(F)| \cdot \|U - W\|_\square$ (Lovász [Lov12, Lemma 10.23]).

Theorem 16 (Counting lemma)**.**

Proof. □

Chapter 6

Compactness

The cut-distance quotient modulo weak isomorphism is a compact metric space (Lovász [Lov12, Theorem 9.23]). Concretely, we prove total boundedness (via the regularity lemma and grid quantization) and completeness (via a direct limit construction from rapidly converging subsequences) of the graphon pseudometric space.

6.1 Total Boundedness

For any $\varepsilon > 0$, there exists a finite ε-net in cut distance. The construction uses the regularity lemma to approximate any graphon by a step graphon, then quantizes coefficients to a grid (Lovász [Lov12, Section 9.3]).

Theorem 17 (Total boundedness of graphon space)**.**

Proof. □

6.2 Completeness

Every Cauchy sequence in cut distance has a limit graphon. The proof extracts a rapidly converging subsequence, builds the limit graphon via Radon–Nikodym, and shows the full sequence converges (Lovász [Lov12, Section 9.3]).

Theorem 18 (Completeness of graphon space)**.**

Proof. □

Chapter 7

Inverse Counting Lemma and Convergence

The inverse counting lemma is the converse of the counting lemma: if all homomorphism densities are similar, then the graphons are close in cut distance. Together with the counting lemma, this gives the fundamental convergence equivalence (Lovász [Lov12, Section 10.6 and Theorem 11.5]).

7.1 Inverse Counting Lemma

For any $\varepsilon > 0$, there exist $\delta > 0$ and a finite family of test graphs such that δ-close homomorphism densities imply ε-close cut distance (Lovász [Lov12, Lemma 10.32]). *Status:* proved modulo both temporary `sorry`s (Rokhlin for partition alignment, algebraic determination for the step graphon core).

Theorem 19 (Inverse counting lemma)**.**

Proof. □

7.2 Convergence Equivalence

A sequence of graphons converges in cut distance if and only if all homomorphism densities converge (Lovász [Lov12, Theorem 11.5]).

Theorem 20 (Convergence equivalence)**.**

Proof. □

Bibliography

[BCL$^+$08] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi, *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Advances in Mathematics **219** (2008), no. 6, 1801–1851.

[FK99] Alan Frieze and Ravi Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), no. 2, 175–220.

[Lov12] László Lovász, *Large networks and graph limits*, Colloquium Publications, vol. 60, American Mathematical Society, 2012.