Hyperfinite graphings and combinatorial optimization

László Lovász
Hungarian Academy of Sciences and Eötvös Loránd University
Budapest
August 10, 2018

Contents

1 Introduction 2
  1.1 Pushing forward and pulling back . . . . . . . . . . . . . . . . . . . . 2
  1.2 Graphings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
  1.3 Hyperfinite graphings . . . . . . . . . . . . . . . . . . . . . . . . . . . 3
  1.4 Convergent graph sequences . . . . . . . . . . . . . . . . . . . . . . . .4

2 Combinatorial version 5
  2.1 Graph partitioning . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
  2.2 Generalization to hypergraphs . . . . . . . . . . . . . . . . . . . . . . .7
  2.3 Linear programming duality . . . . . . . . . . . . . . . . . . . . . . . . 11

3 Fractional partitions in graphings 11
  3.1 The space of connected subgraphs . . . . . . . . . . . . . . . . . . . . . 11
  3.2 Separation in graphings . . . . . . . . . . . . . . . . . . . . . . . . . . 13

4 Duality 15
  4.1 Compact graphings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
  4.2 Duality for graphings . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
  4.3 Completing the proof of Theorem 11 . . . . . . . . . . . . . . . . . . . 21
  4.4 Splitting into finite parts . . . . . . . . . . . . . . . . . . . . . . . . . . 22

5 Open problems 24

Abstract

We exhibit an analogy between the problem of pushing forward measurable sets under measure preserving maps and linear relaxations in combinatorial optimization. We show how invariance of hyperfiniteness of graphings under local isomorphism can be reformulated as an infinite version of a natural combinatorial optimization problem, and how one can prove it by extending well-known
proof techniques (linear relaxation, greedy algorithm, linear programming duality) from the finite case to the infinite. We develop a procedure of compactifying graphings, which may be of interest on its own.

1 Introduction

1.1 Pushing forward and pulling back

Let $X$ and $Y$ be standard probability spaces, and let $\phi : X \to Y$ be a measure preserving map. By definition, we can pull back a measurable subset $Y' \subseteq Y$: the set $\phi^{-1}(Y')$ is measurable in $X$ and has the same measure as $Y'$. It is a lot more troublesome to push forward a measurable subset $X' \subseteq X$: the image $\phi(X')$ may not be measurable, and its measure may certainly differ from the measure of $X'$.

On the other hand, if we have a measure $\mu$ on $X$ (possibly different from the probability measure of $X$), then we can push it forward by the formula $\mu(\phi(Y')) = \mu(\phi^{-1}(Y'))$.

So if we really have to push forward a subset of $X$, we would like to “approximate” it in a suitable sense by a measure, and push forward this measure to $Y$. Then, of course, we still face the task to “distill” a subset of $Y$ from this measure on $Y$.

This vague description may sound familiar to a basic technique in combinatorial optimization: linear relaxation. (It is also similar to “fuzzy sets”, which is a related setup.) The goal of this paper is to show that this analogy is in fact much more relevant than it seems. We show how some important results in graph limit theory (like invariance of hyperfiniteness under local isomorphism), can be reformulated as infinite versions of natural combinatorial optimization problems, and how one can prove them by extending well-known proof techniques from the finite case to the infinite. (Since the proof for the infinite case is described in [10], we only sketch it here.)

1.2 Graphings

For the rest of this paper, we fix a positive integer $D$, and all graphs we consider are supposed to have maximum degree at most $D$.

A graphing is a Borel graph with bounded (finite) degree on a standard probability space $(I, A, \lambda)$, satisfying the following measure-preserving condition for any two Borel subsets $A, B \subseteq I$:

$$\int_A \deg_B(x) d\lambda(x) = \int_B \deg_A(x) d\lambda(x). \quad (1)$$

Here $\deg_B(x)$ denotes the number of edges connecting $x \in I$ to points of $B$. (It can be shown that this is a Borel function of $x$.) Most of the time, we may assume that $I = [0, 1]$ and $\lambda$ is the Lebesgue measure.
Such a graphing defines a measure on Borel subsets of $I^2$: on rectangles we define
\[ \eta(A \times B) = \int_A \deg_B(x) \, d\lambda(x), \]
which extends to Borel subsets in the standard way. We call this the *edge measure* of the graphing. It is concentrated on the set of edges, and it is symmetric in the sense that interchanging the two coordinates does not change it.

For a graphing $G$ and positive integer $r$, let us pick a random element $x \in I$ according to $\lambda$, and consider the subgraph $B_G(x, r)$ induced by nodes at distance at most $r$ from $x$. This gives us a probability distribution on $r$-balls: rooted graphs with degrees at most $D$ and radius (maximum distance from the root) at most $r$. Let $\rho_{G,r}$ be the distribution of $B_G(v, r)$. We sometimes suppress the subscript $G$ when the underlying graphing is understood.

Two graphings $G_1$ and $G_2$ are called *locally equivalent*, if $\rho_{G_1,r} = \rho_{G_2,r}$ for every $r \geq 0$. To characterize local equivalence, let us define a map $\phi: V(G_1) \to V(G_2)$ to be a *local isomorphism* from $G_1$ to $G_2$, if it is measure preserving, and for every $x \in V(G_1)$, $\phi$ is an isomorphism between the connected component of $G_1$ containing $x$ and the connected component of $G_2$ containing $\phi(x)$. It is easy to see that if there exists a local isomorphism $G_1 \to G_2$, then $G_1$ and $G_2$ are locally equivalent.

A local isomorphism may not be bijective, or even injective: it may map different components of $G_1$ on the same component of $G_2$. So it is not sufficient to characterize local equivalence. But making it symmetric, we get a characterization [10]:

**Proposition 1** Two graphings $G_1$ and $G_2$ are locally equivalent if and only if there exists a third graphing $G$ having local isomorphisms $G \to G_1$ and $G \to G_2$.

### 1.3 Hyperfinite graphings

There is an important special class of very “slim” graphings. For a graphing $G$, a set $T$ of edges will be called *$k$-splitting*, if every connected component of $G \setminus T$ has at most $k$ nodes. We denote by $\text{sep}_k(G)$ the infimum of $\eta(T)$, where $T$ is a $k$-splitting Borel set of edges. A graphing $G$ is *hyperfinite*, if $\text{sep}_k(G) \to 0$ as $k \to \infty$.

It is surprising that the following basic fact about hyperfiniteness is nontrivial.

**Theorem 2** Let $G_1$ and $G_2$ be locally equivalent graphings. If $G_1$ is hyperfinite, then so is $G_2$.

This theorem was first proved in [10]. It is closely related to a theorem of Schramm [11] about hyperfiniteness of locally convergent graph sequences; we’ll come back to this connection in the last section. Independently, Elek [5] derived this result from a theorem of Kaimanovich [8]. The proof in [10] yields an explicit relationship between the values $\text{sep}_k(G_1)$ and $\text{sep}_k(G_2)$.
Note that the stronger statement that \( \text{sep}_k(G_1) = \text{sep}_k(G_2) \) is not true (Example 21.12 in [10]). An analogous statement, however, will be true for the measure version (see Theorem 11 below).

To illustrate the “push forward – pull back” problem discussed in the introduction, let us start to prove this theorem. Let \( \varepsilon > 0 \); we want to prove that there is a \( k \geq 1 \) such that \( \text{sep}_k(G_2) < \varepsilon \); in other words, there is a \( k \)-splitting Borel set \( T_2 \subseteq E(G_2) \) such that \( \eta_2(T_2) < \varepsilon \) (where \( \eta_i \) denotes the edge measure of \( G_i \)). By definition, there is a \( k \geq 1 \) and a \( k \)-splitting Borel set \( T_1 \) for \( G_1 \) with \( \eta_1(T_1) < \varepsilon \). By Proposition 1, there is a third graphing \( G \) having local isomorphisms \( \phi_1 : G \to G_1 \) and \( \phi_2 : G \to G_2 \). We can pull back the set \( T_1 \) to \( G \): the set \( T = \phi_1^{-1}(T_1) \) satisfies \( \eta(T) = \eta_1(T_1) < \varepsilon \), and (since \( \phi_1 \) is a local isomorphism from \( G \setminus T \) to \( G_1 \setminus T_1 \)) the connected components of \( G \setminus T \) have no more than \( k \) nodes. This shows that \( \text{sep}_k(G) < \varepsilon \).

To complete the proof, we would like to “push forward” the set \( T \) to \( G_2 \); but we have no control over what happens to its measure. To get around this difficulty, we introduce a fractional version of the \( k \)-splitting problem, which is defined in terms of a measure, and thus it can be pushed forward in a manageable way. But we lose by this, and the main step will be to estimate the loss.

### 1.4 Convergent graph sequences

Graphings were introduced (at least in this setting) as limit objects of locally convergent sequences of bounded degree graphs [1, 4]. Let us sketch this connection.

The probability distribution \( \rho_{G,r} \) on \( r \)-balls can be defined for finite graphs just as for graphings. We say that a sequence \( (G_n : n = 1, 2, \ldots) \) of finite graphs is locally convergent, if for every \( r \geq 1 \), the probability distributions \( \rho_{G_n,r} \) converge (note that these distributions are defined on the same finite set of \( r \)-balls, independently of \( G_n \)). This notion of convergence was introduced by Benjamini and Schramm [2].

We say that \( G_n \to G \) (where \( G \) is a graphing), if \( \rho_{G_n,r} \to \rho_{G,r} \) for every \( r \geq 1 \). For every locally convergent graph sequence \( (G_n : n = 1, 2, \ldots) \) there is a graphing \( G \) such that \( G_n \to G \). This fact can be derived from the work of Benjamini and Schramm; it was stated explicitly in [1] and [4]. It is clear that a convergent graph sequence determines its limit up to local equivalence only.

**Remark 3** Benjamini and Schramm describe a limit object in the form of a probability distribution on rooted countable graphs with degrees at most \( D \), with a certain “unimodularity” condition. This limit object is unique. Graphings contain more information than what is passed on to the limit. Among others, they can represent limits of sequences that are convergent in a stronger sense called local-global convergence [7]. But for us exactly the weaker notion of convergence, and the uncertainty in the limit object it introduces (local equivalence), is interesting.

For a finite graph, the definition of hyperfiniteness makes no sense (every finite graph is hyperfinite); we have to move to infinite families of graphs. A family of finite
graphs is hyperfinite, if for every $\varepsilon > 0$ there is an integer $k \geq 1$ such that every graph $G$ in the family satisfies $\text{sep}_k(G) \leq \varepsilon$. Many important families of graphs (with a fixed degree bound) are hyperfinite: trees, planar graphs, and more generally, every non-trivial minor-closed family \cite{3}. As a non-hyperfinite family, let us mention any expander sequence.

The connection between hyperfinite graph families and hyperfinite graphings is nice, and as it turns out, nontrivial:

**Theorem 4** Let $(G_n : n = 1, 2, \ldots)$ be a sequence of finite graphs with all degrees bounded by $D$, locally converging to a graphing $G$. Then $G$ is hyperfinite if and only if the family $\{G_n : n = 1, 2, \ldots\}$ is hyperfinite.

This theorem is due to Schramm \cite{11} (with a somewhat sketchy proof). A complete proof based on other methods was described by Benjamin, Schramm and Shapira \cite{3}. A third proof could be based on the graph partitioning algorithm of Hassidim, Kelner, Nguyen and Onak \cite{6}. The proof in \cite{10}, which is based on Theorem 2 above, is perhaps closest to Schramm’s original method, although cast in a different form.

## 2 Combinatorial version

### 2.1 Graph partitioning

In this section, we discuss the finite version of the main tool in the proof of Theorem 2. Let $G = (V, E)$ be a finite graph on $n$ nodes. The notion of $k$-splitting edge sets can be defined for $G$ just as for graphings. We denote by $\text{sep}_k(G)$ the minimum of $|T|/n$, where $T$ is a $k$-splitting set of edges.

We formulate a relaxation of the problem of computing $\text{sep}_k(G)$, which can be expressed as a linear program. There are many ways to do so (as usual); we choose one which is perhaps not the simplest, but which will generalize to graphings easily.

Let $\mathcal{R}_k = \mathcal{R}_k(G)$ denote the set of subsets $A \subseteq V$ with $1 \leq |A| \leq k$ that induce a connected subgraph of $G$. An $\mathcal{R}_k$-partition of $V$ is a disjoint subfamily $\mathcal{F} \subseteq \mathcal{R}_k$ covering every node. For $A \subseteq V$, let $\partial A$ denote the set of edges connecting $A$ to $V \setminus A$. We can express $\text{sep}_k(G)$ as

$$\text{sep}_k(G) = \min_{\mathcal{F}} \frac{1}{2n} \sum_{A \in \mathcal{F}} |\partial A|,$$

where $\mathcal{F}$ ranges over all $\mathcal{R}_k$-partitions. Indeed, if $T$ is $k$-splitting, then the components of $G \setminus T$ form an $\mathcal{R}_k$-partition $\mathcal{F}$ with $|T| = \frac{1}{2} \sum_{A \in \mathcal{F}} |\partial A|$. Conversely, for every $\mathcal{R}_k$-partition $\mathcal{F}$, the set $T = \bigcup_{A \in \mathcal{F}} \partial A$ is $k$-splitting, and $\sum_{A \in \mathcal{F}} |\partial A| = 2|T|$ (since every edge in $T$ is counted with two sets $A$).
This suggests the following relaxation: A weighting \( x : R_k \to \mathbb{R} \) is a fractional \( R_k \)-partition, if
\[
x_A \geq 0, \quad \sum_{A \in R_k} x_A = 1 \quad (\forall v \in V).
\]
We define
\[
sep_k^*(G) = \min_x \frac{1}{2n} \sum_{A \in R_k} x_A |\partial A|,
\]
where \( x \) ranges over all fractional \( R_k \)-partitions. The indicator function of an \( R_k \)-partition is a fractional \( R_k \)-partition, and hence
\[
sep_k^*(G) \leq sep_k(G).
\]
Equality does not hold in general: the triangle has \( sep_2(K_3) = 2/3 \) but \( sep_2^*(K_3) = \frac{1}{2} \).
But we have the following weak converse:

**Theorem 5** For every finite graph \( G \) with maximum degree \( D \),
\[
sep_k(G) \leq sep_k^*(G) \left( 2 + \ln \frac{D}{2sep_k^*(G)} \right).
\]
This is the finite version of a result for graphings (Lemma 21.10 in [10]). As we will see, it is crucial that the upper bound depends on \( sep_k^*(G) \) and \( D \) only, not on \( k \) or \( n \).

We give the proof of this theorem in the next section, in a more general form. To get this more general form, we modify the \( k \)-partition problem by looking for covering subgraph-families rather than partitions. This does not change the value of \( sep_k \). More exactly,
\[
sep_k(G) = \min_{F} \frac{1}{2n} \sum_{A \in F} |\partial A|,
\]
where \( F \) ranges over families \( F \subseteq R_k \) covering every node. Indeed, allowing more families \( F \) could only lower the minimum. On the other hand, consider a covering family \( F \) minimizing (6). If this consists of disjoint sets, we are done. Suppose not, and let \( A_1, A_2 \in F \) such that \( A_1 \cap A_2 \neq \emptyset \). Let \( e_1 \) denote the number of edges between \( A_1 \setminus A_2 \) and \( A_1 \cap A_2 \), and define \( e_2 \) similarly. We may assume that \( e_1 \leq e_2 \). Replacing \( A_1 \) by \( A_1' = A_1 \setminus A_2 \), we still have a covering family, and since \( |\partial A_1'| \leq |\partial A_1| + e_1 - e_2 \leq |\partial A_2| \), we have another optimizer in (6). We can repeat this until all sets in \( F \) will be disjoint.

This suggests that we could use another linear relaxation. We define a fractional cover as a vector \( x \subseteq R_k^+ \) such that \( \sum_{A \ni v} x_A \geq 1 \) for all \( v \in V \). We define
\[
sep_k^{**}(G) = \min_x \frac{1}{2n} \sum_{A \in R_k} x_A |\partial A|,
\]
where $x$ ranges over all fractional covers. It is clear that

$$ \text{sep}_k^*(G) \leq \text{sep}_k^*(G), $$

and I don’t know whether strict inequality can ever hold here. Otherwise, there is not much difference in the behavior of these two relaxations. In particular, Theorem 5 would remain valid with $\text{sep}_k^*$ instead of $\text{sep}_k^*$. We are going to use whichever is more convenient.

### 2.2 Generalization to hypergraphs

The theorem can be generalized to hypergraphs. Let $\mathcal{H}$ be a hypergraph on node set $V$, without isolated nodes, and let $w : \mathcal{H} \to \mathbb{R}_+$ be an edge-weighting. An edge-cover is a subset $\mathcal{F} \subseteq \mathcal{H}$, such that $\cup \mathcal{F} = V$. A fractional edge-cover is an assignment of weights $X : \mathcal{H} \to \mathbb{R}_+$ such that

$$ \sum_{A \ni v} x_A \geq 1 \quad (\forall v \in V). \quad (8) $$

Define

$$ \sigma = \sigma(\mathcal{H}, w) = \min_{\mathcal{F}} \frac{1}{n} \sum_{A \in \mathcal{F}} w(A), \quad (9) $$

where $\mathcal{F}$ ranges over all edge-covers, and

$$ \sigma^* = \sigma^*(\mathcal{H}, w) = \min_{x} \frac{1}{n} \sum_{A \in \mathcal{H}} w(A)x_A, \quad (10) $$

where $x$ ranges over all fractional edge-covers. Note that the covering by singletons is in the competition, hence $\sigma^* \leq 1$.

**Theorem 6** Let $\mathcal{H}$ be a hypergraph on node set $V$ with $|V| = n$, such that $\{x\} \in \mathcal{H}$ for each $x \in V$. Let $w : \mathcal{H} \to \mathbb{R}_+$ be an edge-weighting such that $w(\{v\}) \leq 1$ for every $v \in V$. Then

$$ \sigma^* \leq \sigma \leq \sigma^* \left(2 + \ln \frac{1}{\sigma^*}\right). $$

**Proof.** The inequality $\sigma^* \leq \sigma$ is trivial.

To prove the upper bound on $\sigma$, we use a version of the greedy algorithm. For $i = 1, 2, \ldots$, select edges $Y_1, Y_2, \ldots, Y_m \in \mathcal{H}$ so that $Y_i$ is a minimizer of

$$ \min_Y \frac{w(Y)}{|Y \setminus (Y_1 \cup \cdots \cup Y_{i-1})|}. $$

(We don’t consider edges $Y$ for which the denominator is zero.) We stop when $\cup_i Y_i = V$. Let $\mathcal{F} = \{Y_1, \ldots, Y_m\}$. 

Set $y_i = |Y_i \setminus (Y_1 \cup \cdots \cup Y_{i-1})|$ and $w_i = w(Y_i)$. We start with some simple inequalities. First, a partitioning into singletons is a possibility in the definition of $\sigma$, and hence

$$\sigma^* \leq \sigma \leq \frac{1}{n} \sum_{v \in V} w(\{v\}) \leq 1. \quad (11)$$

When choosing $Y_i$ ($i < m$), any edge $A \in H$ with $A \setminus (Y_1 \cup \cdots \cup Y_{i-1}) \neq \emptyset$ was also available, but not chosen. Hence

$$\frac{w_i}{y_i} = \frac{w(Y_i)}{|Y_i \setminus (Y_1 \cup \cdots \cup Y_{i-1})|} \leq \frac{w(A)}{|A \setminus (Y_1 \cup \cdots \cup Y_{i-1})|} \quad (12)$$

In particular,

$$\frac{w_i}{y_i} \leq \frac{w(Y_{i+1})}{|Y_{i+1} \setminus (Y_1 \cup \cdots \cup Y_{i-1})|} \leq \frac{w(Y_{i+1})}{|Y_{i+1} \setminus (Y_1 \cup \cdots \cup Y_i)|} = \frac{w_{i+1}}{y_{i+1}},$$

and so

$$\frac{w_1}{y_1} \leq \frac{w_2}{y_2} \leq \cdots \leq \frac{w_m}{y_m} \quad (13)$$

When choosing $Y_m$, any singleton $v \in V \setminus (Y_1 \cup \cdots \cup Y_{m-1})$ was available, showing that

$$\frac{w_m}{y_m} \leq \frac{w(\{v\})}{1} \leq 1. \quad (14)$$

We claim that

$$y_i + y_{i+1} + \cdots + y_m \leq \frac{y_i}{w_i} \sigma^* n. \quad (15)$$

Indeed, if $x$ is the minimizer in the definition of $\sigma^*$, then using $(12)$,

$$y_i + y_{i+1} + \cdots + y_m = |V \setminus (Y_1 \cup \cdots \cup Y_{i-1})| \leq \sum_{v \in V \setminus (Y_1 \cup \cdots \cup Y_{i-1})} \sum_{A \ni v} x_A = \sum_{A \in H} x_A |A \setminus (Y_1 \cup \cdots \cup Y_{i-1})|$$

$$\leq \sum_{A \in H} x_A \frac{y_i}{w_i} w(A) = \frac{y_i}{w_i} \sigma^* n.$$

Now we turn to bounding $w(F) = w_1 + \cdots + w_m$. Let $1 \leq a \leq m$ be an integer such that

$$\frac{w_{a-1}}{y_{a-1}} < \sigma^* \leq \frac{w_a}{y_a}$$

(possibly $a = 0$ or $a = m + 1$). Then

$$w_1 + \cdots + w_{a-1} \leq \sigma^* (y_1 + \cdots + y_{a-1}).$$
For the remaining sum, multiply \( (15) \) by \( w_a/y_a \) for \( i = a \) and by \( w_i/y_i - w_{i-1}/y_{i-1} \) for \( i > a \), and sum the inequalities. On the left side, the coefficient of \( y_i \) will be

\[
\frac{w_a}{y_a} + \left( \frac{w_{a+1}}{y_{a+1}} - \frac{w_a}{y_a} \right) + \cdots + \left( \frac{w_i}{y_i} - \frac{w_{i-1}}{y_{i-1}} \right) = \frac{w_i}{y_i},
\]

and so on the left side we get

\[
\frac{w_a}{y_a} y_a + \cdots + \frac{w_m}{y_m} y_m = w_a + \cdots + w_m.
\]

On the right side, the coefficient of \( \sigma^* n \) can be estimated as follows, using that \( w_a/y_a \geq \sigma^* \) and \( w_m/y_m \leq 1 \):

\[
\frac{w_a}{y_a} y_a + \sum_{i=a+1}^m \left( \frac{w_i}{y_i} - \frac{w_{i-1}}{y_{i-1}} \right) y_i \leq 1 + \sum_{i=a+1}^m \int \frac{1}{t} \, dt
\]

\[
= 1 + \int_{w_a/y_a}^{w_m/y_m} \frac{1}{t} \, dt \leq 1 + \int_{\sigma^*}^{1} \frac{1}{t} \, dt = 1 + \ln \frac{1}{\sigma^*}.
\]

Thus

\[
w_1 + \cdots + w_m \leq \sigma^* (y_1 + \cdots + y_{a-1}) + \left( 1 + \ln \frac{1}{\sigma^*} \right) \sigma^* n \leq 2 + \ln \frac{1}{\sigma^*} \sigma^* n.
\]

Dividing by \( n \), we get the upper bound in the theorem. \( \square \)

**Remark 7** A simpler but non-algorithmic proof of Theorem 3 can be sketched as follows. Let again \( x \) be the minimizer in the definition of \( \sigma^* \), let \( t > 0 \), and form a family \( \mathcal{F} \) of edges by selecting \( A \in \mathcal{H} \) with probability \( p_A = \min(1, tx_A) \), and adding all singletons that have not been covered. The probability that a node \( v \) is not covered by the randomly selected edges is at most \( e^{-t} \). This is trivial if \( p_A = 1 \) for any of the hyperedges containing \( v \), and else it follows by the estimate

\[
\prod_{A \ni v} (1 - p_A) \leq \exp \left( - \sum_{A \ni v} p_A \right) = \exp \left( - t \sum_{A \ni v} x_A \right) \leq e^{-t}.
\]

So the expected number of edges on the boundaries of sets in \( \mathcal{F} \) is at most

\[
\sum_{A \in \mathcal{H}} tx_A |\partial A| + e^{-t} n = tn\sigma^* + e^{-t} n.
\]

This bound is minimized when \( t = - \ln \sigma^* \), giving a very little better bound than in the theorem.

The first proof we gave has two advantages: first, it is algorithmic, and second (perhaps more importantly from our point of view), it generalizes to the case of graphings.
Remark 8 If all weights are 1, and all edges \( A \in \mathcal{H} \) have cardinality \(|A| \leq k\), then it is easy to see that \( \sigma^* \geq 1/k \), and hence
\[
\sigma \leq \sigma^* \left(2 + \ln \frac{1}{\sigma^*}\right) \leq \sigma^* \left(2 + \ln k\right).
\]
With some care, we could reduce the first term from 2 to 1. This is a well known inequality [9].

Remark 9 If we only know that \(|A| \leq k\) for all \( A \in \mathcal{H} \), but make no assumption about the weights, we can still prove the (rather trivial) inequality
\[
\sigma \leq k \sigma^*.
\] (16)
Indeed, for every node \( v \), let \( Z_v \in \mathcal{H} \) be an edge containing \( v \) with minimum weight. Then \( \{Z_v : v \in V\} \) is an edge-cover, and hence
\[
\sigma \leq \sum_v w(Z_v).
\]
On the other hand, let \( (x_A : A \in \mathcal{H}) \) be an optimal fractional edge-cover, then \( \sum_{A \ni v} x_A \geq 1 \), and hence
\[
\sum_v w(Z_v) \leq \sum_v w(Z_v) \sum_{A \ni v} x_A = \sum_A x_A \sum_{v \in A} w(Z_v)
\leq \sum_A x_A \sum_{v \in A} w(A) = \sum_A x_A |A| w(A) \leq k \sum_A x_A w(A) = k \sigma^*.
\]

Proof of Theorem 5 We consider the hypergraph \((V, \mathcal{R}_k(G))\) of connected subgraphs of size at most \( k \), with edge-weights \( w(Y) = |\partial Y|/D \). We claim that
\[
\text{sep}_k(G) = \frac{D}{2} \sigma(\mathcal{H}, w).
\] (17)
Indeed, for any family \( F \subseteq \mathcal{R}_k \), we have
\[
\frac{1}{2} \sum_{A \in F} |\partial A| = \frac{D}{2n} \sum_{A \in F} w(A),
\]
and the quantities on both sides of (17) are the minima of the two sides of this last equation. The inequality
\[
\text{sep}_k^*(G) \geq \text{sep}_k^* (G) = \frac{D}{2} \sigma^*(\mathcal{H}, w)
\] (18)
follows similarly. Using (17) and (18), we can apply Theorem 6
\[
\text{sep}_k(G) = \frac{D}{2} \sigma \leq \frac{D}{2} \sigma^* \left(2 + \ln \frac{1}{\sigma^*}\right) \leq \text{sep}_k^* \left(2 + \ln \frac{D}{2 \text{sep}_k^* (G)}\right).
\]
2.3 Linear programming duality

To motivate our results about duality for separation in graphings, let us describe the (very standard) formulation of the dual of the fractional $k$-separation problem.

The definition of $\text{sep}_k^*(G)$ can be considered as a linear program, with a variable $x_A$ for each set $A \in \mathcal{R}_k$:

\[
\begin{align*}
\text{minimize} & \quad \sum_{A \in \mathcal{R}_k} \frac{|\partial A|}{2n} x_A, \\
\text{subject to} & \quad x_A \geq 0, \\
& \quad \sum_{A \ni v} x_A \geq 1 \quad (\forall v \in V). \\
\end{align*}
\]

If $k$ is bounded, then this program has polynomial size, so together with (5) and Theorem 5, we get a polynomial time approximation algorithm for $\text{sep}_k^*(G)$, with a multiplicative error of $O(\ln(1/\text{sep}_k^*(G)))$.

The dual program has a variable $y_v$ for each node $v$, and has the following form (after some scaling):

\[
\begin{align*}
\text{maximize} & \quad \frac{1}{2n} \sum_{v \in V} y_v, \\
\text{subject to} & \quad y_v \geq 0, \\
& \quad \sum_{v \in A} y_v \leq |\partial A| \quad (\forall A \in \mathcal{R}_k). \\
\end{align*}
\]

This will give us a hint how to formulate "dual" in the graphing setting.

3 Fractional partitions in graphings

3.1 The space of connected subgraphs

We consider a graphing $G = (I, A, \lambda, E)$, and fix an integer $k \geq 1$. Just as for graphs, $\mathcal{R}_k = \mathcal{R}_k(G)$ is the set of subsets of $I$ inducing a connected subgraph with at most $k$ nodes. Every singleton set belongs to $\mathcal{R}_k$. Note that $\mathcal{R}_k$ can be thought of as a subset of $I \cup I^2 \cup \cdots \cup I^k$; it is easy to see that this is a Borel subset. We can introduce a metric $d_k$ on $I \cup I^2 \cup \cdots \cup I^k$: we use the $\ell_\infty$ metric on each $I^r$, and distance 1 between points in different sets $I^r$. It is clear that $\mathcal{R}_k$ is a closed subset of $I \cup I^2 \cup \cdots \cup I^k$, and so it is a compact space.

Let $\tau$ be a finite measure on $\mathcal{R}_k$. We define the marginal of $\tau$ by

\[
\tau'(X) = \int_{\mathcal{R}_k} \frac{|X \cap Y|}{|Y|} d\tau(Y) \quad (X \in \mathcal{A}).
\]
In the case when $\tau$ is a probability measure, the probability distribution $\tau'$ could be generated by selecting a set $Y \in \mathcal{R}_k$ according to $\tau$, and then a point $y \in Y$ uniformly. For every function $g : I \to \mathbb{R}$ and every finite set $Y \subseteq I$, define

$$\bar{g}(Y) = \frac{1}{|Y|} \sum_{y \in Y} g(y).$$

Then for every finite measure $\tau$ on $\mathcal{R}_k$, and every integrable function $g : I \to \mathbb{R}$, we have

$$\int_{\mathcal{R}_k} \bar{g} \, d\tau = \int_I g \, d\tau'.$$  \(21\)

It is easy to see that this equation characterizes $\tau'$.

Mainly for technical purposes, we define a specific probability distribution $\mu$ on the Borel sets of $\mathcal{R}_k$ by selecting a random point $x \in I$ according to $\lambda$, and then a subset $Y \ni x$ inducing a connected subgraph, uniformly among all such subsets. Since there are only a finite number of such subgraphs, and at least one, this makes sense. If $\mathcal{R}_k(x)$ denotes the set of sets $Y \in \mathcal{R}_k$ containing $x$, then

$$\mu(S) = \int_I \frac{|S \cap \mathcal{R}_k(x)|}{|\mathcal{R}_k(x)|} \, d\lambda(x).$$

An obvious but important property of $\mu$ is that if $\mu(S) = 0$ for some $S \subseteq \mathcal{R}_k$, then $\lambda(\cup S) = 0$.

For a Borel subsets $T, U \subseteq \mathcal{R}_k$, define

$$\Phi_T(x) = |T \cap \mathcal{R}_k(x)| \quad \text{and} \quad \mu_T(Z) = \mu(T \cap U).$$  \(22\)

We need the following identity:

$$\frac{d\mu_T'}{d\lambda}(x) = \Phi_T(x).$$  \(23\)

(The left side is defined for almost all $x \in I$ only.) To prove this, let $X$ be a Borel subset of $I$, and define

$$f(x, y) = \begin{cases} \sum_{Y : x, y \in Y \in T} \frac{1}{|Y|} & \text{if } x \in X, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly $f(x, y) \neq 0$ implies that $x$ and $y$ belong to the same component of $G$ and their distance is at most $k - 1$. Thus the sums in the following equation are finite, and we can apply the Mass Transport Principle (see e.g. Theorem 18.49 in [11]):

$$\int_I \sum_y f(x, y) \, d\lambda(x) = \int_I \sum_x f(x, y) \, d\lambda(y).$$
On the left, we get
\[
\int \sum_I f(x, y) \, d\lambda(x) = \int \sum_I \sum_{Y \ni x, y \in T} \frac{1}{|Y|} \, d\lambda(x) = \int \sum_I \sum_{Y \ni x \in T} \frac{|Y|}{|Y|} \, d\lambda(x) = \int \Phi_T \, d\lambda,
\]
while on the right,
\[
\int \sum_I f(x, y) \, dy = \int \sum_I \sum_{x \in X} \sum_{y \ni x \in Y \in T} \frac{1}{|Y|} \, dy = \int \sum_I \sum_{y \ni x \in Y \in T} \frac{|X \cap Y|}{|Y|} \, dy = \mu_T'(X).
\]
This proves (23).

Let \( g : I \rightarrow \mathbb{R} \) be a bounded Borel function and define the integral measure
\[
\lambda_g(X) = \int_X g \, d\lambda.
\]
We need the identity
\[
\int_I \Phi_T \, d\lambda \, \mu_T = \int_I \Phi_T(x) \, d\lambda_g. \tag{24}
\]
Indeed, using (23),
\[
\int_{\mathcal{R}_k} g \, d\mu_T = \int_I g \, d\mu_T' = \int_I g \, \frac{d\mu_T'}{d\lambda} \, d\lambda = \int_I g \Phi_T \, d\lambda = \int_I \Phi_T(x) \, d\lambda_g.
\]

3.2 Separation in graphings

The definition \( \text{sep}^*_k \) can be extended to graphings with some care. Here the probabilistic version of the definition is easier to use. We say that \( \tau \) is a fractional \( \mathcal{R}_k \)-partition, if its marginal is the uniform distribution on \( I \), and \( \tau \) is a fractional \( \mathcal{R}_k \)-cover, if its marginal majorizes the uniform distribution.

For a finite set \( Y \subseteq V(G) \), we define its edge expansion (or isoperimetric number) by
\[
h(Y) = \frac{|\partial Y|}{|Y|} \tag{25}
\]
We define the relaxed \( k \)-splitting value in terms of the “expected expansion” of a fractional partition \( \tau \); more exactly,
\[
\text{sep}^*_k(G) = \frac{1}{2} \inf_{\tau} \mathbb{E} h(Y) = \frac{1}{2} \inf_{\mathcal{R}_k} \int h(Y) \, d\tau(Y), \tag{26}
\]

13
where $\tau$ ranges over all fractional $R_k$-partitions, and $Y$ is a random subset from the distribution $\tau$.

Just as in the finite case, we would get the same value if we allowed $R_k$-covers instead of $R_k$-partitions.

The main theorem relating $k$-splitting numbers and their fractional versions extends to graphings:

**Theorem 10** For every graphing $G$ with maximum degree $D$,

$$\text{sep}_k^*(G) \leq \text{sep}_k(G) \leq \text{sep}_k^*(G) \log \frac{8D}{\text{sep}_k^*(G)}.$$ 

As an immediate corollary, we see that a graphing $G$ is hyperfinite if and only if $\text{sep}_k^*(G) \to 0$ as $k \to \infty$.

The main difficulty in extending the previous proof to graphings is that the greedy selection of the sets $Y_j$, as described in the previous section, would last through an uncountable number of steps, thus messing up all measurability conditions. Instead, we do the selection in phases, where a (typically) uncountable number of disjoint approximate minimizers are selected simultaneously. It turns out that one can stop after a finite number of such phases. The prize we have to pay is a small loss in the constant, as the logarithm here is binary. (This construction uses some results from the theory of Borel graphs.) We refer to [10] for details.

The following theorem shows that $\text{sep}_k^*$ behaves better that $\text{sep}_k$ with respect to local equivalence.

**Theorem 11** Let $G_1$ and $G_2$ be locally equivalent graphings. Then for every $k \geq 1$, $\text{sep}_k^*(G_1) = \text{sep}_k^*(G_2)$.

Let us describe again the trivial half of the proof. The nontrivial half will be given in Section 4.3. It will be based expressing $\text{sep}_k^*(G)$ in a dual form, in terms of the existence of a function on $V(G)$, and on the fact that (with some care) functions can be pushed in both directions along a measure preserving map.

By Proposition 1, it suffices to prove it in the case when there is a local isomorphism $\phi : G_1 \to G_2$. There is an easy direction: There is a fractional $R_k(G_1)$-partition $\tau_1$ such that for a random set $Y \in R_k(G_1)$ from this distribution, $E(|\partial Y|/|Y|) \leq \text{sep}_k(G_1)$.

Since $\phi(Y) \in R_k(G_2)$ for every $Y \in R_k(G_1)$ by the definition of local isomorphism, $\phi$ defines a map $\tilde{\phi} : R_k(G) \to R_k(G_2)$. It is easy to see that this is a measurable map, and hence it pushes forward the measure $\tau_1$ to a probability measure $\tau_2$ on $R_k(G_2)$. Furthermore, $\tau_2$ is a fractional partition (this follows since $\phi_2$ is measure preserving), and

$$\int_{R_k(G_2)} h(Y) d\tau_2(Y) = \int_{R_k(G_1)} h(Y) d\tau_1(Y).$$
Thus $\text{sep}_k^*(G_2) \leq \text{sep}_k^*(G_1)$.

The previous two theorems imply a quantitative version of Theorem 2:

**Corollary 12** Let $G_1$ and $G_2$ be locally equivalent graphings. Then

$$\text{sep}_k(G_2) \leq \text{sep}_k(G_1) \log \frac{8D}{\text{sep}_k(G_1)}.$$ 

Indeed, we have $\text{sep}_k^*(G_2) = \text{sep}_k^*(G_1) \leq \text{sep}_k^*(G_2)$, and so by Theorem 10,

$$\text{sep}_k(G_2) \leq \text{sep}_k^*(G_2) \log \frac{8D}{\text{sep}_k^*(G_2)} \leq \text{sep}_k(G_1) \log \frac{8D}{\text{sep}_k(G_1)}$$

(using that the function $x \log(8D/x)$ is monotone increasing for $x \leq 1$).

Note that we have applied the nontrivial half of Theorem 10, but (with some care) only the trivial half of Theorem 11.

**Remark 13** The inequality (16), along with its proof, generalizes to graphings with no difficulty: For every graphing $G$,

$$\text{sep}_k(G) \leq k \text{sep}_k^*(G).$$

However, this inequality would not be good enough to prove Theorem 2. Going through the proof, one could realize the significance that the upper bound in Theorem 10 does not depend on $k$: we need the fact that if $\text{sep}_k(G_1) \to 0$ as $k \to \infty$, then $\text{sep}_k(G_1) \log \frac{8D}{\text{sep}_k(G_1)} \to 0$, and hence $\text{sep}_k(G_2) \to 0$.

## 4 Duality

Our goal is to generalize the linear programming duality of the separation problem to the graphing case. Before developing this, we need a technical construction to compactify graphings.

### 4.1 Compact graphings

Let $G$ be a graphing. For two points $x, y \in V(G)$, we define a *neighborhood isomorphism* between $x$ and $y$ as an isomorphism $\phi: B(x, r) \to B(y, r)$ for some $r \geq 0$ such that $\phi(x) = y$. The number $r$ is the *radius* if the neighborhood isomorphism. Such a neighborhood isomorphism always exists with radius 0.

We say that a graphing $G$ is *compact*, if $V(G) = J$ is a compact metric space, $E$ is a closed subset of $J \times J$, and for every $\varepsilon > 0$ there is a $\delta = \delta(G, r) > 0$ such that if $d(x, y) < \delta$ ($x, y \in J$), then there is a neighborhood isomorphism $\phi$ between $x$ and $y$ with radius $r \geq 1/\varepsilon - 1$ such that $d(z, \phi(z)) \leq \varepsilon$ for every $z \in B(x, r)$.

We say that a graphing $G = (I, \mathcal{A}, \lambda, E)$ is a *strong subgraphing* of a graphing $G' = (I', \mathcal{A}', \lambda', E')$, if $G$ is the union of connected components of $G'$, the Borel
subsets of $I$ are the restrictions of the Borel subsets of $I'$ to $V(G)$, and $\lambda'(X) = \lambda(I \cap X)$ for every Borel subset of $I'$. Since we assume that the probability spaces underlying our graphings are standard, it follows that $I$ is measurable in $I'$ in the Lebesgue sense, and its measure is 1. It follows that $I$ contains a Borel subset $I^*$ with $\lambda(I^*) = \lambda'(I^*) = 1$.

It is easy to see that if $G$ is a strong subgraphing of $G'$, then $\text{sep}_k(G) = \text{sep}_k(G')$ and $\text{sep}_k^*(G) = \text{sep}_k^*(G')$. In particular, $G$ is hyperfinite if and only if $G'$ is. (In fact, $G$ and $G'$ are local-global equivalent in the sense of [7], and hyperfiniteness is invariant under local-global equivalence; we don’t define this somewhat complicated notion here.)

**Theorem 14** Every graphing is a strong subgraphing of a compact graphing.

The compact graphing constructed below will be called a compactification of the original graphing.

**Proof.** Let us start with any compact metric $d_0$ on $I = V(G)$ giving the prescribed Borel sets; for example, we can identify $I$ with $[0, 1]$ with the euclidean metric.

For two points $x, y \in V(G)$ and a neighborhood isomorphism $\phi$ between $x$ and $y$ with radius $r$, we define

$$|\phi| = \max\left\{\frac{1}{r + 1}, \max_{z \in \phi(B(x, r))} d_0(z, \phi(z))\right\} \quad\text{and}\quad d(x, y) = \inf_{\phi} |\phi|,$$

(27)

where $\phi$ ranges over all neighborhood isomorphisms. It is easy to see that $d$ is a metric, and

$$d_0(x, y) \leq d(x, y) \leq 1.$$

Let $(J, d)$ be the completion of the space $(I, d)$.

In what follows, we denote (as before) by $B(x, r)$ the ball in the graphing $G$ (here $r$ is always a nonnegative integer), and by $\mathcal{V}(x, \varepsilon)$, an open ball in $(J, d)$ with center $x$ and radius $\varepsilon$ ($0 < \varepsilon \leq 1$).

**Claim 1** The Borel sets in $(I, d_0)$ are exactly the restrictions to $I$ of Borel sets in $(J, d)$.

The fact that $d_0 \leq d$ implies that if a set $U$ is open in $(I, d_0)$, then it is open in $(I, d)$, and so it is the restriction of an open set of $(J, d)$ to $I$. Restrictions of Borel sets of $(J, d)$ to $I$ form a sigma-algebra, which contains all open sets of $(I, d_0)$, and hence, all Borel sets.

The converse is a bit more elaborate. It suffices to show that the restriction to $I$ of a ball $\mathcal{V} = \mathcal{V}(x, \varepsilon)$ of $(J, d)$ is a Borel set in $(I, d_0)$ (not necessarily open).

First, suppose that $x \in I$. Then $y \in \mathcal{V}$ if and only if there exists a neighborhood isomorphism $\phi$ from $x$ to $y$ with $|\phi| \leq \varepsilon$. Let $\mathcal{B}_r$ be the set of points $y \in \mathcal{V} \cap I$
for which such a neighborhood isomorphism exists with radius $r$. Clearly $r < 1/\varepsilon$ if $\mathcal{B}_r \neq \emptyset$.

Let $B(x, r) = \{x = z_1, z_2, \ldots, z_N\}$. Consider the set $S$ of points $(y_1, \ldots, y_N) \in I^N$ for which

$$y_i y_j \in E(G) \iff x_i x_j \in E(G) \quad \text{and} \quad d_0(x_i, y_i) \leq \varepsilon \quad (1 \leq i < j \leq N).$$

It is clear that $S$ is Borel in $(I, d_0)^N$. The projection of $S$ to the first coordinate is just the set $\mathcal{B}_r$. If $(y_1, \ldots, y_N) \in S$, then the points $y_1, \ldots, y_N$ induce a connected subgraph of $G$, and so every point $y_1$ is contained in a finite number of such $N$-tuples. Thus Lusin’s Theorem implies that the projection of $S$ to the first coordinate is Borel.

It follows that $\mathcal{B} \cap I = \bigcup_r \mathcal{B}_r$ is Borel in $(I, d_0)$. Second, if $x \in J \setminus I$, then let $(x_1, x_2, \ldots)$ be a sequence of points in $I$ such that $x_n \to x$. The identity

$$\mathcal{B} = \bigcup_{k \geq 1} \bigcap_{n \geq 1} \bigcup_{m \geq n} \mathcal{B}\left(x_m, \varepsilon - \frac{1}{k}\right),$$

shows that $\mathcal{B} \cap I$ is Borel in $(I, d_0)$. So every open ball, restricted to $I$, is a Borel set in $I$.

Since open balls generate the sigma-algebra of Borel sets, it follows that for every Borel set $U$ in $(J, d)$, the intersection $U \cap I$ is Borel in $(I, d_0)$. This completes the proof of the claim.

We define a graphing $\hat{G}$ on $J$. The underlying probability measure $\hat{\lambda}$ can be defined on the Borel subsets of $(J, d)$ by

$$\hat{\lambda}(X) = \lambda(X \cap I)$$

(since the set $X \cap I$ is Borel in $(I, d_0)$). We define $E(\hat{G})$ as the closure of $E(G)$ in $(J, d)^2$. This guarantees that it is a symmetric Borel set in $(J, d)$. It is easy to check that $\hat{\lambda}$ satisfies the identity \([I]\), so $\hat{G}$ is a graphing.

We show that it is a compact graphing. This will follow from the following two assertions.

**Claim 2** $(J, d)$ is compact.

It suffices to prove that every infinite sequence of points $(x_n \in I : n = 1, 2, \ldots)$ has a subsequence that is Cauchy. Since there are only a bounded number of possible neighborhoods of any radius, we can select a subsequence such that $B(x_n, r)$ is isomorphic to $B(x_r, r)$ for $n \geq r$. We fix neighborhood isomorphisms $\phi_r : B(x_r, r) \to B(x_{r+1}, r)$, and let $\phi_{r,n} = \phi_{n-1} \circ \cdots \circ \phi_r$ ($n \geq r$).

Going to a subsequence again, we may assume that for every $r \geq 0$ and every $z \in B(x_r, r)$, the sequence $(\phi_{r,n}(z) : n = r, r+1, \ldots)$ is convergent in the $d_0$ metric. Since $B(x_r, r)$ is finite, for every $\varepsilon > 0$ there is an $N = N(\varepsilon, r)$ such that for $n, m > N$ we have $d_0(\phi_{r,n}(z), \phi_{r,m}(z)) < \varepsilon$. 17
Now let $\varepsilon > 0$ and $r = \lceil 1/\varepsilon \rceil$. For $n > m > \max(r, N(\varepsilon, r)$, let $\phi$ denote the restriction of $\phi_{n-1} \circ \cdots \circ \phi_m$ to $\phi_{r,m}(B(x, r))$. Then $\phi$ is neighborhood isomorphism between $x_m$ and $x_n$, and

$$|\phi| = \max\left\{ \frac{1}{r+1}, \max_{z \in \phi(B(x, r))} d_0(z, \phi(z)) \right\} \leq \varepsilon.$$  

This shows that $d(x_m, x_n) < \varepsilon$, which proves the claim.

Claim 3 For $x, y \in J$ there is a neighborhood isomorphism $\phi$ of $\hat{G}$ between $x$ and $y$ with radius $r \geq 1/(2d(x, y)) - 1$ such that $d(z, \phi(z)) \leq 2d(x, y)$ for every $z \in B(x, r)$.

Note that with $d_0(z, \phi(z)) \leq d(x, y)$ instead of $d(z, \phi(z)) \leq 2d(x, y)$ this would be trivial. By the definition of $d$, there is a neighborhood isomorphism $\phi$ between $x$ and $y$ with radius $r_0 \geq 1/d(x, y) - 1$, such that

$$d_0(z, \phi(z)) \leq d(x, y)$$

(28)

for all $z \in B(x, r_0)$. Let $r = \lfloor r_0/2 \rfloor$. For every $z \in B(x, r)$, we have $B(z, r) \subseteq B(x, r_0)$, and so restricting $\phi$ to $B(z, r)$ we get a neighborhood isomorphism $\phi_z$ between $z$ and $\phi(z)$ with radius $r$. Thus by (28),

$$d(z, \phi(z)) \leq |\phi_z| \leq \max\left\{ \frac{1}{r+1}, d(x, y) \right\} \leq 2d(x, y).$$

This proves the claim. Thus $\hat{G}$ is indeed a compact graphing.

To prove that $G$ is a strong subgraphing of $\hat{G}$, all that is left to show is that $G$ is the union of components of $\hat{G}$. In other words:

Claim 4 For $x \in I$ and $y \in J$, we have $xy \in \hat{E}$ if and only if $xy \in E(G)$.

The “if” part is trivial. Let $xy \in \hat{E}$, then there are edges $x_ny_n \in E(G)$ such that $x_n \to x$ and $y_n \to y$. There are neighborhood isomorphisms $\phi_n$ from $x_n$ to $x$ such that $|\phi_n| \to 0$. This implies that the corresponding radii $r_n \to \infty$, and hence $y_n = \phi_n(y_n)$ exists for sufficiently large $n$ and $xy_n' \in E(G)$. Since $x$ has a finite number of neighbors, we can select an infinite subsequence for which $y_n' = y'$ is independent of $n$. Let $\phi_n'$ be the restriction of $\phi_n$ to $B(y_n, r_n - 1)$. Then simple computation shows that $|\phi_n'| \leq 2|\phi_n| \to 0$, and so $d(y_n, y') \to 0$. But $d(y_n, y') \to 0$ by hypothesis, and hence $y = y'$ and $xy = xy' \in E(G)$. This proves the claim.

We may do another “purifying” operation of $\hat{G}$, by deleting all points outside the (closed) support of the measure $\lambda$. The following example shows how this eliminates some minor complications.

Example 1 The following example is illustrative. Consider the graphing $C_{\alpha}$ defined on the unit circle by connecting two points if their angular distance is $2\alpha \pi$. Then
every connected component of $C_\alpha$ is a 2-way infinite path. It is clear that $C_\alpha$ is a compact graphing, with the metric $d_0$ equal to the angular metric.

Let $C'_\alpha$ be obtained by deleting an edge $uv$ from $C_\alpha$. There will be two components that are one-way infinite paths $P_1$ and $P_2$ starting at $u$ and $v$; the rest is unchanged. The graphing $C'_\alpha$ is not compact with the same metric, since the 1-ball about $u$ (an edge) is not isomorphic to the 1-ball about any nearby point (two edges). The metric $d$, as constructed above, remains the angular distance between any two points not on $P_1 \cup P_2$; the distance of points on each $P_i$ from points outside $P_1 \cup P_2$ will be positive (but decreasing as we go along the path). We can think of lifting these countably many points off the cycle.

Compactifying $C_\alpha$ fills in the positions of the original points, and the edges between them. So we are left with the compact graphing $C_\alpha$, with two one-way infinite paths spiralling closer and closer to it. The support of the measure $\hat{\lambda}$ is just the original circle. So the purified compactification of $C'_\alpha$ is just $C_\alpha$.

4.2 Duality for graphings

Recall that fractional partitions are probability distributions $\tau$ on $\mathcal{R}^k$ such that $\tau' = \lambda$. This can be expressed explicitly by the conditions

$$\int_{\mathcal{R}^k} \mathcal{G} \, d\tau = \int_{0}^{1} g \, d\lambda \quad (29)$$

for all integrable functions $g : I \to \mathbb{R}_+$. If the graphing $G$ is compact, then it suffices to require (29) for continuous functions $g$. The relaxed separation number $\text{sep}_k^*$ is defined by

$$\text{sep}_k^*(G) = \frac{1}{2} \inf \int_{\mathcal{R}^k} h \, d\tau. \quad (30)$$

A dual characterization (as a supremum instead of an infimum) generalizing (21), is given in the following theorem.

**Theorem 15** For every graphing $G = (I, \mathcal{A}, \lambda, E)$, we have

$$\text{sep}_k^*(G) = \sup \frac{1}{2} \int_{I} g \, d\lambda,$$

where $g$ ranges over all bounded Borel functions $g : I \to \mathbb{R}$ such that $\mathcal{G} \leq h$ on $\mathcal{R}^k$.

Note that the condition on $g$ can be written as $\sum_{y \in Y} g(y) \leq |\partial Y|$ for all $Y \in \mathcal{R}^k$. The proof will show that if the graphing is compact, then we can let $g$ range over continuous functions only.
Proof. We may assume that the graphing is compact; else, we go to its compactification, find the appropriate functions \( g \), and restrict them to the original points, where they are still bounded Borel functions.

Let us denote the supremum on the right side by \( \alpha_k \). First, we show the “easy direction”. For every measure \( \tau \) on \( \mathcal{R}_k \) with \( \tau' = \lambda \), and every Borel function \( g \) as in the theorem, we have

\[
\int_I g \, d\lambda = \int_I g \, d\tau' = \int_{\mathcal{R}_k} \overline{g} \, d\tau \leq \int_{\mathcal{R}_k} h \, d\tau,
\]

proving that \( \text{sep}_k^* (G) \geq \alpha_k \).

To prove the opposite inequality, fix any \( s < 2 \text{sep}_k^* (G) \). Let \( C(I) \) be the space of continuous functions on \( I \), then for every \( g \in C(I) \), we have \( \overline{g} \in C(\mathcal{R}_k) \). Define

\[
K_0 = \left\{ g \in C(I) : \int_I g \, d\lambda = s \right\} \quad \text{and} \quad K = \left\{ f \in C(\mathcal{R}_k) : \exists g \in K_0, \overline{g} \leq f \right\}.
\]

Both \( K_0 \) and \( K \) are convex, and \( K \) has nonempty interior. Both \( K_0 \) and \( K \) contain the constant function \( s \) (defined on \( I \) and \( \mathcal{R}_k \), respectively).

Recalling the metric \( d_k \) on \( I \cup I^2 \cup \cdots \cup I^k \), we can state the following.

**Claim 5** The function \( h \) is continuous in the metric \( d_k \).

Indeed, fix a set \( Y \in \mathcal{R}_k \), and choose \( N > k \) so that any two points in \( Y \cup \partial Y \) are at least \( 1/N \) apart. If \( d_k(Y, Z) < \min(1/N, \delta(N)) \), then \( |Y| = |Z| \), and the isomorphism \( \phi : B(y, N) \to B(z, N) \) (where \( y \in Y \), \( z \in Z \) and \( d(y, z) < 1/N \)) maps \( Y \) onto \( Z \) and their neighborhoods onto each other, showing that \( h(Z) = h(Y) \).

The main step in the proof is the following claim.

**Claim 6** \( h \in K \).

Suppose not, then by the Hahn–Banach Theorem, there is a nonzero continuous linear functional \( \ell \) on \( C(\mathcal{R}_k) \) such that \( \ell(f) \geq \ell(h) \) for every \( f \in K \). Since \( \ell \) remains bounded from below on \( K \), it must be nonnegative on nonnegative functions. We may normalize \( \ell \) so that \( \ell(1) = 1 \). For any \( g \in K_0 \), the functional remains bounded from below on the linear variety of functions of the form \( \overline{f} \), where \( f = s + t(g - s) \) for some \( t \in \mathbb{R} \). This implies that it must be constant on this variety and so it must satisfy \( \ell(\overline{g}) = \ell(s) = s \ell(1) = s \). Thus we have \( \ell(h) \leq \ell(s) = s \).

By the Riesz Representation Theorem, we can represent \( \ell \) by a measure \( \tau \) on \( \mathcal{R}_k \) such that \( \ell(f) = \int_{\mathcal{R}_k} f \, d\tau \) for every \( f \in C(\mathcal{R}_k) \). By the normalization \( \ell(1) = 1 \), \( \tau \) is a probability measure. Furthermore,

\[
\ell(\overline{g}) = \int_{\mathcal{R}_k} \overline{g} \, d\tau = \int_I g \, d\tau' = s = \int_I g \, d\lambda
\]
for every $g \in K_0$, which implies that $\tau' = \lambda$. But then

$$\ell(h) = \int_{\mathcal{R}_k} h \ d\tau \geq 2\text{sep}_k^*(G) > s,$$

a contradiction.

So $h \in K$, and hence there is a nonnegative function $g \in K_0$ such that $\tau \leq h$. So $2\alpha_k \geq s$. Since this holds for every $s < 2\text{sep}_k^*(G)$, we must have $\alpha_k = \text{sep}_k^*(G)$.  

4.3 Completing the proof of Theorem 11

We have two graphings $G_t = (I_t, A_t, \lambda_t, E_t) \ (t = 1, 2)$ and a local isomorphism $\phi : G_1 \to G_2$. We have seen that $\text{sep}_k^*(G_2) \leq \text{sep}_k^*(G_1)$. We want to prove that equality holds here.

To this end, let $0 < s < \text{sep}_k^*(G_1)$, and recall the formula for $\text{sep}_k^*(G_1)$ from Theorem 15: there is a bounded Borel function $g_1 : I_1 \to \mathbb{R}_+$ such that

$$\frac{1}{2} \int_{I_1} g_1 \ d\lambda_1 > s.$$

To “push” $g_1$ to $G_2$, consider the measure $\lambda_{g_1}$ defined by

$$\lambda_{g_1}(X) = \int_X g_1 \ d\lambda_1,$$

and its “push-forward” defined by $\gamma(X) = \lambda_{g_1}(\phi^{-1}(X))$. Since $g_1$ is bounded, we have $\lambda_{g_1} \leq K\lambda$ for some $K > 0$, and hence $\gamma(X) \leq \lambda_1(\phi^{-1}(X)) = \lambda_2(X)$. It follows that the Radon–Nikodym derivative

$$g = \frac{d\gamma}{d\lambda_2}$$

exists and it is bounded by $K$.

We claim that $g$ satisfies the conditions for $G_2$ in Theorem 15. The only nontrivial part is to show that $\overline{g} \leq h$.

Let us apply definitions (22) to each $G_t$, to get functions $\Phi_t$ and measures $\mu_t$ and $\mu_t^T$. It is not hard to verify (using that $\phi$ is a local isomorphism) that

$$\Phi_{\phi^{-1}(T)}(x) = \Phi_T^2(\phi(x)) \quad (31)$$

and for all $Y \in \mathcal{R}_k(G_1)$,

$$h(\phi(Y)) = h(Y).$$
Let $T_2 \subseteq \mathcal{R}_k(G_2)$ be any Borel set and let $T_1 = \phi^{-1}(T_2)$. By (24) and (31), we have
\[
\int_{T_2} g \, d\mu_2 = \int_{T_1} \Phi^2_{T_2}(x) \, d\gamma(x) = \int_{T_1} \Phi^1_{T_1}(x) \, d\lambda_1(x) = \int_{T_1} \gamma_1 \, d\mu_1
\]
\[
\leq \int_{T_1} h \, d\lambda_1 = \int_{T_2} h \, d\mu_2
\]
Since this holds for every $T_2$, it follows that $\gamma \leq h$ almost everywhere on $\mathcal{R}_k(G_2)$. Changing the value of $g$ to 0 on all point of all sets $Y \in \mathcal{R}_k(G_2)$ where $g(Y) > h(Y)$, we get a function satisfying the conditions of Theorem 15, and hence
\[
\text{sep}_k^*(G_2) \geq \int_{I_2} g \, d\lambda_2 = \int_{I_2} 1 \, d\gamma = \int_{I_1} g_1 \, d\lambda_1 > s.
\]
Since this holds for every $s < \text{sep}_k^*(G_1)$, it follows that $\text{sep}_k^*(G_2) \geq \text{sep}_k^*(G_1)$.

### 4.4 Splitting into finite parts

It is an alternative definition of hyperfinite graphs that one can delete a set of edges with arbitrarily small measure so that the remaining graph has finite components. If the graph is not hyperfinite, we may be interested in determining how small the measure of such an edge set can be. This question motivates the following discussion.

Clearly $\text{sep}_k$ is monotone decreasing in $k$, and hence $\lim_{k \to \infty} \text{sep}_k(G) = \text{sep}(G)$ exists. It is not hard to see that $\text{sep}_k(G)$ is the infimum of edge-measures of edge sets, whose deletion from $G$ results in a graph with every component finite.

We can define similarly $\lim_{k \to \infty} \text{sep}_k^*(G) = \text{sep}^*(G)$, which is related to $\text{sep}(G)$ by similar inequalities as in Theorem 10:
\[
\text{sep}^*(G) \leq \text{sep}(G) \leq \text{sep}_k^*(G) \log \frac{8D}{\text{sep}^*(G)}.
\]

The graphing $G$ is hyperfinite if and only if $\text{sep}(G) = 0$ or, equivalently, $\text{sep}^*(G) = 0$. Duality gives the following nice formula for $\text{sep}^*(G)$ of any graphing.

**Theorem 16** For every graphing,
\[
\text{sep}^*(G) = \sup_g \frac{1}{2} \int_I g \, d\lambda,
\]
where $g$ ranges over all bounded Borel functions $g : V(G) \to \mathbb{R}_+$ such that
\[
\sum_{y \in Y} g(y) \leq |\partial Y|
\]
for every finite set $Y \subseteq V(G)$ inducing a connected subgraph.
Proof. The proof is similar to that in section 4.3, but the details are different. By Theorem 15, there exist bounded Borel functions $g_k \geq 0$ ($k = 1, 2, \ldots$) such that
\[
\sum_{y \in Y} g_k(y) \leq |\partial Y| \quad (\forall Y \in \mathcal{R}_k),
\]
and
\[
\int_I g_k \, d\lambda \geq \text{sep}^*_k(G) - \frac{1}{k}. \tag{34}
\]
Note that $g_k(x) = \mathcal{F}_k(\{x\}) \leq |\partial \{x\}| \leq D$, so these functions remain uniformly bounded. It follows by Alaoglu’s Theorem that they have a weak* limit $g$ in $L^\infty(I, \lambda)$; this limit satisfies
\[
\int_I g_k f \, d\lambda \to \int_I g f \, d\lambda \quad (k \to \infty) \tag{35}
\]
for every $f \in L^1(I, \lambda)$. In particular, it follows that $0 \leq g \leq D$ (almost everywhere, but we may change $g$ on a zero set, so that this holds everywhere), and
\[
\|g\|_1 = \int_I g \, d\lambda \geq \text{sep}^*(G).
\]
We claim that $\mathcal{F} \leq h$ holds almost everywhere on $\mathcal{R}_1 \cup \mathcal{R}_2 \cup \ldots$. It suffices to prove this on $\mathcal{R}_k$ for a fixed $k$.

Let $T \subseteq \mathcal{R}_k$ be any Borel set. The Radon-Nikodym derivative $\Phi_T = d\mu_T / d\lambda$ exists and is bounded by (23). By (24),
\[
\int_T \mathcal{F} \, d\mu = \int_I g \, d\mu_T = \int_I g \, \Phi_T \, d\lambda.
\]
Similarly,
\[
\int_T g_k \, d\mu = \int_I g_k \, \Phi_T \, d\lambda.
\]
Since $\mathcal{F} \leq h$, we have
\[
\int_I g_k f \, d\lambda \leq \int_T h \, d\mu,
\]
and hence by weak convergence,
\[
\int_T \mathcal{F} \, d\mu \leq \int_T h \, d\mu.
\]
This holds for every $T$, which implies that $\mathcal{F} \leq h$ almost everywhere. Changing $g$ to 0 on all points of $I$ in sets violating $\mathcal{F} \leq h$, we get a function as in the theorem. \qed
Corollary 17 A graphing \( G \) is not hyperfinite if and only if there exists a bounded Borel function \( g : V(G) \to \mathbb{R}_+ \), not almost everywhere zero, such that
\[
\sum_{y \in Y} g(y) \leq |\partial Y|
\]
for every finite set \( Y \subseteq V(G) \) inducing a connected subgraph.

5 Open problems

1. Is the logarithmic factor needed in Theorems 8 and 10, or in inequality (32)? Could a constant factor be enough? For Theorem 6, in the general setting of hypergraphs, it is easy to show that a logarithmic factor is needed. For example, let \( H \) be the hypergraph whose vertices are all \( p \)-element subsets of an \( 2p \)-element set \( S \), and edges are subfamilies containing a given element of \( S \). Let all hyperedges have weight 1. Then any \( p + 1 \) hyperedges cover \( V(H) \), but no \( p \) of them does, and so
\[
\sigma(H, w) = \frac{p + 1}{2p}
\]
On the other hand, \( x_A = 1/p \) (\( A \in E(H) \)) defines a fractional cover, which is easily seen to be optimal, and hence
\[
\sigma^*(H, w) = \frac{2}{2p}.
\]
So we see that
\[
\sigma(H, w) \sim \frac{1}{4} \sigma^*(H, w) \log \frac{1}{\sigma^*(H, w)}
\]

2. Is \( \text{sep}_k^*(G) = \text{sep}_k^*(G) \) for every graph \( G \)? In other words, can the simple manipulation described after the statement of Theorem 8 be extended from covers to fractional covers? If not, how far can these two parameters be?

3. Can we improve the bound in Theorem 8 by allowing larger components? Perhaps it is true that
\[
\lim_{m \to \infty} \text{sep}_m(G) \leq \text{sep}_k^*(G)
\]
for every graphing \( G \) and integer \( k \geq 1 \). This would imply that \( \text{sep}^*(G) = \text{sep}(G) \) for every graphing \( G \).
References

[1] D. Aldous and R. Lyons: Processes on Unimodular Random Networks, *Electron. J. Probab.* **12**, Paper 54 (2007), 1454–1508.

[2] I. Benjamini and O. Schramm: Harmonic functions on planar and almost planar graphs and manifolds, via circle packings. *Invent. Math.* **126** (1996), 565–587.

[3] I. Benjamini, O. Schramm and A. Shapira: Every Minor-Closed Property of Sparse Graphs is Testable, *Advances in Math.* **223** (2010), 2200–2218.

[4] G. Elek: On limits of finite graphs, *Combinatorica* **27** (2007), 503–507.

[5] G. Elek: Finite graphs and amenability, *J. Func. Anal.* **263** (2012), 2593–2614.

[6] A. Hassidim, J.A. Kelner, H.N. Nguyen and K. Onak: Local Graph Partitions for Approximation and Testing, *Proc. 50th Ann. IEEE Symp. on Found. Comp. Science* (2009), 22–31.

[7] H. Hatami, L. Lovász and B. Szegedy: Limits of locally-globally convergent graph sequences, *Geom. Func. Anal.* **24** (2014), 269–296.

[8] V. Kaimanovich: Amenability, hyperfiniteness, and isoperimetric inequalities, *C. R. Acad. Sci. Paris Sr. I Math.* **325** (1997), 999–1004.

[9] L. Lovász: On the ratio of optimal fractional and integral covers, *Discrete Math.* **13** (1975), 383–390.

[10] L. Lovász: *Large networks and graph limits*, Amer. Math. Soc., Providence, R.I. (2012).

[11] O. Schramm: Hyperfinite graph limits, *Elect. Res. Announce. Math. Sci.* **15** (2008), 17–23.