

Nondeterministic graph property testing

László Lovász

Katalin Vesztegombi

Institute of Mathematics, Eötvös Loránd University

Budapest, Hungary

February 2012

Abstract

A property of finite graphs is called nondeterministically testable if it has a “certificate” such that once the certificate is specified, its correctness can be verified by random local testing. In this paper we study certificatees that consist of one or more unary and/or binary relations on the nodes, in the case of dense graphs. Using the theory of graph limits, we prove that nondeterministically testable properties are also deterministically testable.

1 Introduction

Let \mathcal{P} be a property of finite simple graphs (i.e., a class of finite simple graphs closed under isomorphism). We say that \mathcal{P} is *testable*, if there exists another property \mathcal{T} (called a *test property*) satisfying the following conditions:

- if a graph G has property \mathcal{P} , then for all $1 \leq r \leq |V(G)|$, a random induced subgraph on r nodes (chosen uniformly among all such induced subgraphs) has property \mathcal{T} with probability at least $2/3$, and
- for every $\varepsilon > 0$ there is a $r_\varepsilon \geq 1$ such that if G is a graph whose edit distance from \mathcal{P} is at least $\varepsilon|V(G)|^2$, then for all $r_\varepsilon \leq r \leq |V(G)|$, a random induced subgraph on r nodes has property \mathcal{T} with probability at most $1/3$.

This notion of testability is often called *oblivious testing*, which refers to the fact that no information about the size of G is assumed. The definition extends trivially to graphs whose nodes and/or edges are colored with a fixed finite number k of colors.

Let L be a graph whose nodes and edges are k -colored ($k \geq 1$). We call such a graph briefly a *k -colored graph*. It will be convenient to assume that the graph G itself is complete; we can consider the missing edges as having a further color. Given a positive integer $m \leq k$, we can get an ordinary graph from L by keeping only the edges with colors $1, \dots, m$, and then forgetting the coloring. We will denote this graph by L' . If \mathcal{Q} is a property of colored graphs, then we define $\mathcal{Q}' = \{L' : L \in \mathcal{Q}\}$.

A graph property \mathcal{P} is *nondeterministically testable*, if there exist two numbers $k \geq m \geq 1$ and a property \mathcal{Q} of k -colored graphs such that \mathcal{Q} is testable and $\mathcal{Q}' = \mathcal{P}$. In other words, G has property \mathcal{P} if and only if we can color the nodes with k colors,

the edges with m colors and non-edges with further $k - m$ colors so that the resulting k -colored graph has property \mathcal{Q} . We call such a coloring a *certificate* for \mathcal{P} .

Instead of a k -coloring, we could specify k unary and k symmetric binary relations on $V(G)$ as a certificate (this would be more in the spirit of mathematical logic). The fact that in a coloring they are disjoint and partition $V(G)$ and $\binom{V(G)}{2}$, respectively, can be easily tested. Conversely, such a system of relations defines a 2^k -coloring. As long as we are not concerned with efficiency, these two ways of looking at certificates are equivalent.

Clearly every testable property is nondeterministically testable (choosing $k = m = 1$). Our main result asserts the converse.

Theorem 1.1 *A graph property is nondeterministically testable if and only if it is testable.*

The proof uses the theory of graph limits as developed in [4, 8], and its connection with property testing [10].

2 Preliminaries

2.1 Convergence and limits

Convergence of a sequence of dense finite graphs was defined by Borgs, Chayes, Lovász, Sós and Vesztegombi [3, 4]. Graphons were introduced by Lovász and Szegeedy in [8] as limits of convergent sequences of finite graphs. We have to extend these notions to colored graphs; this was essentially done in [11], but we use here a little different terminology.

For a graph G and integer $r \leq |V(G)|$, let us select an ordered r -tuple of nodes of G randomly and uniformly (without repetition). Let $\mathbb{G}(r, G)$ denote the subgraph induced by these nodes. If G is colored, then $\mathbb{G}(r, G)$ is also a colored graph in the obvious way.

We say that a sequence of colored graphs L_n is *convergent*, if $|V(G_n)| \rightarrow \infty$, and for every $r \geq 1$, the distribution of $\mathbb{G}(r, L_n)$ tends to a limit as $n \rightarrow \infty$. Note that this distribution is over a finite set, so it does not matter in which norm we want it to converge.

The limit object of a convergent sequence of simple graphs can be described as a symmetric measurable function $W : [0, 1]^2 \rightarrow [0, 1]$, called a *graphon*. We will need the more general notion of a *kernel*, a bounded symmetric measurable function $W : [0, 1]^2 \rightarrow \mathbb{R}$.

For a sequence of k -colored graphs, the limit object is a bit more complicated: it consists of k graphons (W^1, \dots, W^k) such that $\sum_h W^h = 1$; and also of a partition (coloring) $A^1 \cup \dots \cup A^k$ of $[0, 1]$ into measurable sets. We call the $(2k)$ -tuple $\mathbf{W} = (W^1, \dots, W^k, A^1, \dots, A^k)$ a *k-graphon*.

Let L be a k -colored graph with $V(L) = [n]$. Let \mathcal{J}_n denote the partition of $[0, 1]$ into n intervals J_1, \dots, J_n of equal length. We can associate with L a k -graphon $\mathbf{W}_L = (W_L^1, \dots, W_L^k, A_L^1, \dots, A_L^k)$, where A_L^h is the union of intervals J_i for which i has color h in L , and

$$\mathbf{W}_L^h(x, y) = \begin{cases} 1, & \text{if } x \in J_i, y \in J_j, \text{ and the color of } ij \text{ is } h, \\ 0, & \text{otherwise.} \end{cases}$$

More generally, we can consider a *fractionally k -colored graph* H in which for the nodes are k -colored, and for every unordered pair (i, j) of nodes, we have k nonnegative weights $\beta^1(i, j), \dots, \beta^k(i, j)$ whose sum is 1. (We could also consider fractional node-colorings, but these are not needed to prove our theorem, and they would not lead to greater generality.) We consider every k -colored graph as a special case, where β^i is the indicator function of the i -th edge color. For a fractionally k -colored graph H , we define the k -graphon \mathbf{W}_H in the obvious way.

We can sample a k -colored graph $\mathbb{G}(r, \mathbf{W})$ on node set $[r]$ from a k -graphon \mathbf{W} as follows: we choose r independent random points $X_1, \dots, X_r \in [0, 1]$ uniformly; we color $i \in [r]$ with color h if $X_i \in A_h$, and we color a pair (i, j) with color h with probability $W^h(X_i, X_j)$ (independently for different pairs of nodes).

The following fact was proved in [11].

Proposition 2.1 *Let L_n be a convergent sequence of k -colored graphs. Then there is a k -graphon \mathbf{W} such that $\mathbb{G}(r, L_n) \rightarrow \mathbb{G}(r, \mathbf{W})$ in distribution.*

We write $L_n \rightarrow \mathbf{W}$ if this holds.

Let W be a graphon, and let \mathcal{J} be a partition of $[0, 1]$ into measurable sets J_1, \dots, J_m with positive measure. We denote by $W_{\mathcal{J}}$ the graphon obtained by averaging W in every rectangle $J_i \times J_j$. More precisely, for $x \in J_i$ and $y \in J_j$ we define

$$W_{\mathcal{P}}(x, y) = \frac{1}{\lambda(J_i)\lambda(J_j)} \int_{J_i \times J_j} W(u, z) du dz.$$

We quote a well-known fact (for a proof see e.g. Lemma 3.5 in [4]):

Proposition 2.2 *For every kernel W , we have $W_{\mathcal{J}_n} \rightarrow W$ almost everywhere.*

2.2 Many distances

Convergence to a k -graphon can be described in more explicit forms. Let us start with recalling the *cut distance* of two graphs G and G' on the same node set V (introduced by Frieze and Kannan [7]):

$$d_{\square}(G, G') = \max_{S, T \subseteq V} \frac{e_G(S, T) - e_{G'}(S, T)}{|V|^2},$$

where $e_G(S, T)$ denotes the number of edges with one endpoint in S and the other in T . This can be extended to two fractionally k -colored graphs H and H' on the same node set V , with edgeweights $\beta_H^h(i, j)$ and $\beta_{H'}^h(i, j)$ ($h = 1, \dots, k$), by the hairy (but in fact quite natural) formula

$$d_{\square}(H, H') = \frac{D_1}{|V|} + \sum_{h=1}^k \max_{S, T \subseteq V} \left| \sum_{\substack{i \in S \\ j \in T}} (\beta_H^h(i, j) - \beta_{H'}^h(i, j)) \right|,$$

where D_1 is the number of nodes colored differently in H and H' .

A related notion for kernels is the *cut norm*

$$\|W\|_{\square} = \sup_{S, T \subseteq [0,1]} \int_{S \times T} W(x, y) dx dy.$$

The cut norm has many nice properties (see [4]), of which we need the following (Lemma 2.2 in [10]):

Proposition 2.3 *Let W_n ($n = 1, 2, \dots$) be a sequence of uniformly bounded kernels such that $\|W_n\|_{\square} \rightarrow 0$. Then for every bounded measurable function $Z : [0, 1]^2 \rightarrow \mathbb{R}$, we have $\|W_n Z\|_{\square} \rightarrow 0$.*

(It would be enough to assume that Z is integrable, and the proof in [10] would work for the general case with a little care.)

The cut norm defines a distance function between two kernels by $d_{\square}(U, W) = \|U - W\|_{\square}$. From the point of view of graph limits, however, a kernel is only relevant up to a measure preserving transformation of $[0, 1]$. Hence it is often more natural to consider the following distance notion, which we call the *cut distance*:

$$\delta_{\square}(U, W) = \inf_{\phi, \psi} \|U^{\phi} - W^{\psi}\|_{\square},$$

where $\phi, \psi : [0, 1] \rightarrow [0, 1]$ are measure preserving maps, and $U^{\phi}(x, y) = U(\phi(x), \phi(y))$. This defines a pseudometric on the set of kernels (it is only a pseudometric, since different kernels may have measure 0). An important fact is that endowing the space of all graphons with this pseudometric makes it compact [9].

We generalize the cut norm distance to two k -graphons $\mathbf{U} = (U^1, \dots, U^k, A^1, \dots, A^k)$ and $\mathbf{W} = (W^1, \dots, W^k, B^1, \dots, B^k)$:

$$d_{\square}(\mathbf{U}, \mathbf{W}) = \sum_{h=1}^k \lambda(A^h \triangle B^h) + \sum_{h=1}^k \|U^h - W^h\|_{\square}. \quad (1)$$

Similarly as above, we need the “unlabeled distance”

$$\delta_{\square}(\mathbf{U}, \mathbf{W}) = \inf_{\phi, \psi} d_{\square}(\mathbf{U}^{\phi}, \mathbf{W}^{\psi}) \quad (2)$$

(where \mathbf{U}^{ψ} is obtained from \mathbf{U} by substituting $\phi(x)$ for x in each of the $2k$ functions constituting \mathbf{U}). We can use this notion to define the cut distance of two fractionally k -colored graphs H and H' by

$$\delta_{\square}(H, H') = \delta_{\square}(\mathbf{W}_H, \mathbf{W}_{H'}).$$

It is easy to see that if $V(H) = V(H')$, then

$$\delta_{\square}(H, H') \leq d_{\square}(H, H').$$

The following result is proved (in a more general form) in [11].

Proposition 2.4 *Let L_n be a sequence of k -colored graphs, and let \mathbf{W} be a k -graphon. Then $L_n \rightarrow \mathbf{W}$ if and only if $\delta_{\square}(\mathbf{W}_{L_n}, \mathbf{W}) \rightarrow 0$.*

We cannot claim convergence in d_{\square} distance, since \mathbf{W}_{L_n} depends on the labeling of the nodes of L_n , while the convergence $L_n \rightarrow \mathbf{W}$ does not. However, at least in the case of simple graphs, the following stronger version is true ([4], Theorem 4.16):

Proposition 2.5 *Let G_n be a sequence of graphs, and let W be a graphon such that $G_n \rightarrow W$. Then the graphs G_n can be labeled so that $\|W_{G_n} - W\|_{\square} \rightarrow 0$.*

From the point of view of property testing, the “edit distance” is very important (in fact, from the point of view of analysis, property testing is about the interplay between the edit distance and the cut distance). For two graphs G and G' on the same set of nodes $V = V(G) = V(G')$, their *edit distance* is defined by

$$d_1(G, G') = \frac{|E(G) \Delta E(G')|}{|V|^2}.$$

We generalize this for two k -colored graphs L and L' on the same node set:

$$d_1(L, L') = \frac{D_1}{|V|} + \frac{D_2}{|V|^2},$$

where D_1 is the number of nodes colored differently, and D_2 is the number of edges colored differently, in L and L' .

For two kernels, their edit distance is just their L_1 -distance as functions. For two k -graphons, their edit distance is defined by a formula very similar to (1):

$$d_1(\mathbf{U}, \mathbf{W}) = \sum_{h=1}^k \lambda(A^h \Delta B^h) + \sum_{h=1}^k \|U^h - W^h\|_1.$$

Similarly to (2), we could define the unlabeled version of the edit distance, but we don't need it in this paper.

The following (easy) characterization of testability of graph properties was formulated in [10], Theorem 3.20.

Proposition 2.6 *A graph property \mathcal{P} is testable if and only if for any two sequences (G_n) and (G'_n) of graphs such that $|V(G_n)|, |V(G'_n)| \rightarrow \infty$, $\delta_{\square}(G_n, G'_n) \rightarrow 0$ and $G'_n \in \mathcal{P}$, we have $d_1(G_n, \mathcal{P}) \rightarrow 0$.*

3 Main proof

We start with a randomized construction to obtain a k -colored graph from a fractionally k -colored graph H . We keep the color of every node. We color every edge $ij \in \binom{[n]}{2}$ with color h with probability $\beta^h(i, j)$. For different pairs i, j we make an independent decision. We denote this random k -colored graph by $\mathbb{L}(H)$.

Lemma 3.1 *Let H be a fractionally k -colored graph on n nodes. Then*

$$d_{\square}(H, \mathbb{L}(H)) \leq \frac{10k}{\sqrt{n}}$$

with probability at least $1 - ke^{-n}$.

Proof. For two edge colors, this is just Lemma 4.3 in [4]. For general k , it follows by applying this fact to each edge-color separately. \square

The main step in the proof of Theorem 1.1 is the following lemma.

Lemma 3.2 *Let $\mathbf{W} = (W^1, \dots, W^k, B^1, \dots, B^k)$ be a k -graphon, and let $U = \sum_{h=1}^m W^h$ (where $1 \leq m \leq k$). Let G_n be a sequence of graphs such that $G_n \rightarrow U$. Then there exist k -colored graphs L_n on $V(G_n)$ such that $L'_n = G_n$ and $L_n \rightarrow \mathbf{W}$.*

Proof. First, we construct a fractionally k -colored graph H_n on $V_n = V(G_n)$. To keep the notation simple, assume that $V_n = [n]$. We may assume without loss of generality that B_1, \dots, B_k are intervals of $[0, 1]$. By Proposition 2.5, we can choose the labeling of the nodes of each G_n so that $\|W_{G_n} - U\|_{\square} \rightarrow 0$. We color node i with color h if the left endpoint of the interval J_i belongs to interval B^h .

For every pair $i, j \in [n]$, we select any $x \in J_i$ and $y \in J_j$, and define

$$\beta^h(i, j) = \begin{cases} A_{ij} \frac{W_{\mathcal{J}_n}^h(x, y)}{U_{\mathcal{J}_n}(x, y)} & \text{if } 1 \leq h \leq m, \\ (1 - A_{ij}) \frac{W_{\mathcal{J}_n}^h(x, y)}{1 - U_{\mathcal{J}_n}(x, y)} & \text{if } m + 1 \leq h \leq k, \end{cases}$$

where A_n denotes the adjacency matrix of G_n (these numbers are independent of the choice of x and y). It is easy to check that $\sum_h \beta^h(i, j) = 1$ for all $i \neq j$. We show that for the fractionally k -colored graph H_n constructed this way, we have

$$d_{\square}(\mathbf{W}_{H_n}, \mathbf{W}) \rightarrow 0. \quad (3)$$

Let C^h denote the union of those intervals J_i for which i got color h , i.e., whose left endpoint is contained in B^h . Then we have

$$d_{\square}(\mathbf{W}_{H_n}, \mathbf{W}) = \sum_{h=0}^k \lambda(B^h \triangle C^h) + \|W_{H_n}^h - W^h\|_{\square}.$$

It suffices to prove the convergence for every fixed h . Clearly C^h is an interval, and the symmetric difference $B^h \triangle C^h$ is covered by two intervals J_i . So $\lambda(B^h \triangle C^h) \leq 2/n$, and each term in the first sum tends to 0.

The proof for the second term is trickier. We describe it for $h \leq m$; the other case is analogous. Since $0 \leq W^h \leq U$, we can write $W^h = UZ$, where $0 \leq Z \leq 1$, and $Z = 0$ if $U = 0$. Then we have

$$\|W_{H_n}^h - W^h\|_{\square} = \sup_{S, T \subseteq [0, 1]} \left| \int_{S \times T} W_{H_n}^h - W^h \right|.$$

Substituting from the definition,

$$\int_{S \times T} (W_{H_n}^h - W^h) = \int_{S \times T} \left(W_{G_n}(x, y) \frac{W_{\mathcal{J}_n}^h(x, y)}{U_{\mathcal{J}_n}(x, y)} - W^h(x, y) \right) dx dy.$$

We split this integral as follows:

$$\int_{\substack{S \times T \\ U=0}} W_{G_n} \frac{(UZ)_{\mathcal{J}_n}}{U_{\mathcal{J}_n}} + \int_{\substack{S \times T \\ U \neq 0}} W_{G_n} \left(\frac{(UZ)_{\mathcal{J}_n}}{U_{\mathcal{J}_n}} - Z \right) + \int_{S \times T} (W_{G_n} - U)Z. \quad (4)$$

The first term, which is nonnegative, can be estimated as follows:

$$\int_{\substack{S \times T \\ U=0}} W_{G_n} \frac{(UZ)_{\mathcal{J}_n}}{U_{\mathcal{J}_n}} \leq \int_{S \times T} W_{G_n} \mathbb{1}_{U=0} = \int_{S \times T} (W_{G_n} - U) \mathbb{1}_{U=0} \leq \|(W_{G_n} - U) \mathbb{1}_{U=0}\|_{\square}.$$

Here the right hand side tends to 0 by Proposition 2.3. The second term can be estimated like this. By Proposition 2.2, we have $(UZ)_{\mathcal{J}_n} \rightarrow UZ$ and $U_{\mathcal{J}_n} \rightarrow U$ almost everywhere. Hence $(UZ)_{\mathcal{J}_n}/U_{\mathcal{J}_n} \rightarrow Z$ in almost every point where $U \neq 0$. Since the integrand is bounded, this implies that

$$\int_{\substack{S \times T \\ U \neq 0}} W_{G_n} \left(\frac{(UZ)_{\mathcal{J}_n}}{U_{\mathcal{J}_n}} - Z \right) \rightarrow 0.$$

Finally for the third term in (4), we have

$$\left| \int_{S \times T} (W_{G_n} - U)Z \right| \leq \|(W_{G_n} - U)Z\|_{\square},$$

and here the right hand side tends to 0, again by Proposition 2.3. This proves (3).

To complete the proof of the lemma, we consider the graphs $L_n = \mathbb{L}(H_n)$. By Lemma 3.1, we have

$$d_{\square}(L_n, H_n) \leq \frac{20k}{\sqrt{n}} \quad (5)$$

with probability at least $1 - e^{-n}$. Since $\sum_n e^{-n}$ is convergent, the Borel–Cantelli Lemma implies that almost surely (5) holds for all but a finite number of indices n . Choosing the L_n so that this occurs, we have $d_{\square}(L_n, H_n) = d_{\square}(\mathbf{W}_{L_n}, \mathbf{W}_{H_n}) \rightarrow 0$, and hence $d_{\square}(\mathbf{W}_{L_n}, \mathbf{W}) \rightarrow 0$. \square

Proof of Theorem 1.1. Let \mathcal{P} be a nondeterministically testable property; we show that it is testable. By Proposition 2.6 it suffices to prove that if (G_n) is a sequence of graphs such that $d_{\square}(G_n, \mathcal{P}) \rightarrow 0$, then $d_1(G_n, \mathcal{P}) \rightarrow 0$.

Since \mathcal{P} is nondeterministically testable, there are integers $1 \leq m \leq k$ and a testable property \mathcal{Q} of k -colored graphs such that $\mathcal{P} = \mathcal{Q}'$. Let $\hat{G}_n \in \mathcal{P}$ such that $d_{\square}(G_n, \hat{G}_n) \rightarrow 0$. Since $\hat{G}_n \in \mathcal{P}$, there are k -colored graphs $\hat{L}_n \in \mathcal{Q}$ such that $\hat{G}_n = \hat{L}'_n$. By selecting a subsequence, we may assume that the sequence (\hat{L}_n) is convergent. Let \mathbf{W} be a k -graphon representing its limit, and let $U = \sum_{h=1}^m W^h$. Then $\hat{G}_n \rightarrow U$. From $\delta_{\square}(G_n, \hat{G}_n) \rightarrow 0$ we see that $G_n \rightarrow U$.

Now we invoke Lemma 3.2, and construct k -colored graphs L_n such that $L'_n = G_n$ and $L_n \rightarrow \mathbf{W}$. Hence $d_{\square}(L_n, \mathcal{Q}) \rightarrow 0$. Since \mathcal{Q} is testable, this implies that $d_1(L_n, \mathcal{Q}) \rightarrow 0$, and so we can change the color of $o(n)$ nodes and $o(n^2)$ edges in L_n so that the resulting k -colored graph M_n belongs to \mathcal{Q} . But then $M'_n \in \mathcal{P}$, and M'_n differs from G_n in $o(n^2)$ edges only, so $d_1(G_n, \mathcal{P}) \leq d_1(G_n, M'_n) \rightarrow 0$. \square

4 Applications

One of the first nontrivial results about property testing concerned the maximum cut. Let us derive one version. The property of 2-node-colored graphs that “at least 99% of all edges connect nodes with different colors” is trivially testable, and hence:

Corollary 4.1 *The property of a graph that its maximum cut contains at least 99% of the edges is testable.*

In [10], the *upward closure* of a graph property \mathcal{P} was defined as the graph property \mathcal{P}^\uparrow consisting of those graphs that have a spanning subgraph in \mathcal{P} . Suppose that \mathcal{P} is testable, then the property of 3-edge-colored graphs that “edges with color 1 form a graph with property \mathcal{P} ” is testable, and hence:

Corollary 4.2 *The upward closure of a testable graph property is testable.*

Suppose again that \mathcal{P} is testable, then the property of 4-edge-colored graphs that “edges with colors 1 and 2 form a graph with property \mathcal{P} , and edges with colors 2 and 3 are fewer than 1% of all edges with colors 1 and 3” is testable, and hence:

Corollary 4.3 *If \mathcal{P} is a testable property, then the property that “we can change 1% of the edges to get a graph with property \mathcal{P} ” is also testable.*

5 Concluding remarks

There are several possible analogues and extensions of our results.

One could consider certificates in the form of non-symmetric binary relations. For example, the property of a graph that it has a transitive orientation can be certified by such an orientation. In this case, one could allow that the original graph is directed. The arguments above extend to this case without much difficulty.

A more substantial extension would be to allow certificates in the form of k -ary relations for any k . One could then allow hypergraphs instead of the original graphs. A limit theory for hypergraphs is available (Elek and Szegedy [6], and we expect our main result to generalize to hypergraphs; however, several of the auxiliary results we have made use of have not been extended, and a full proof will take further research.

A generalization in a different direction would be to consider, instead of coloring, node and edge decorations from a compact topological space. For example, the property of being a threshold graph can be certified by a decoration of the nodes by numbers from $[0, 1]$. The limit theory for graphs has been extended to compact decorations [11]; perhaps our main result extends too, but this takes further research.

Finally, let us mention that the situation is quite different in the case of graphs with bounded degree (for which a limit theory analogous to the dense case is available, and property testing has been extensively studied). Here the sampling method is to select r random nodes (uniformly), and explore their neighborhoods of depth r . The property of a graph G that “ G is the disjoint union of two graphs on at least $|V(G)|/3$ nodes” can be certified by coloring the nodes in these two graphs with different colors, so this property is nondeterministically testable. On the other hand, sampling will not distinguish between an expander graph and the disjoint union of two copies of it, so this property is not testable.

Acknowledgements. Research was supported by the European Research Council Grant No. 227701 and by the National Science Foundation under agreement No. DMS-0835373. Hospitality of the Institute for Advanced Study is gratefully acknowledged.

Any opinions and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or of the ERC.

References

- [1] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy: Efficient testing of large graphs, *Combinatorica* **20** (2000) 451–476.
- [2] N. Alon and A. Shapira: A Characterization of the (natural) Graph Properties Testable with One-Sided Error, *SIAM J. Computing* **37** (2008), 1703–1727.
- [3] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi: Counting graph homomorphisms, in: *Topics in Discrete Mathematics* (ed. M. Klazar, J. Kratochvíl, M. Loeb, J. Matoušek, R. Thomas, P. Valtr), Springer (2006), 315–371.
- [4] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi: Convergent Graph Sequences I: Subgraph frequencies, metric properties, and testing, *Advances in Math.* **219** (2008), 1801–1851.
- [5] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi: Convergent Graph Sequences II: Multiway Cuts and Statistical Physics, *Annals of Math.* <http://www.cs.elte.hu/~lovasz/ConvRight.pdf>
- [6] G. Elek, B. Szegedy: Limits of hypergraphs, removal and regularity lemmas. A non-standard approach. <http://arxiv.org/0705.2179>
- [7] A. Frieze and R. Kannan: Quick approximation to matrices and applications, *Combinatorica* **19** (1999), 175–220.
- [8] L. Lovász, B. Szegedy: Limits of dense graph sequences, *J. Comb. Theory B* **96** (2006), 933–957.
- [9] L. Lovász, B. Szegedy: Szemerédi’s Lemma for the analyst, *Geom. Func. Anal.* **17** (2007), 252–270.
- [10] L. Lovász, B. Szegedy: Testing properties of graphs and functions, *Israel J. Math.* **178** (2010), 113–156.
- [11] L. Lovász, B. Szegedy: Limits of compact decorated graphs <http://arxiv.org/abs/1010.5155>