



Accumulation points of the edit distance function

Christopher Cox^{*,1}, Ryan R. Martin², Daniel McGinnis³

Department of Mathematics, Iowa State University, Ames, IA, USA

ARTICLE INFO

Article history:

Received 24 July 2021

Received in revised form 5 February 2022

Accepted 12 February 2022

Available online xxxx

Keywords:

Edit distance

Colored regularity graphs

ABSTRACT

Given a hereditary property $\mathcal{H}$ of graphs and some $p \in [0, 1]$, the edit distance function $\text{ed}_{\mathcal{H}}(p)$ is (asymptotically) the maximum proportion of “edits” (edge-additions plus edge-deletions) necessary to transform any graph of density p into a member of $\mathcal{H}$. For any fixed $p \in [0, 1]$, $\text{ed}_{\mathcal{H}}(p)$ can be computed from an object known as a colored regularity graph (CRG). This paper is concerned with those points $p \in [0, 1]$ for which infinitely many CRGs are required to compute $\text{ed}_{\mathcal{H}}$ on any open interval containing p ; such a p is called an accumulation point. We show that, as expected, $p = 0$ and $p = 1$ are indeed accumulation points for some hereditary properties; we additionally determine the slope of $\text{ed}_{\mathcal{H}}$ at these two extreme points. Unexpectedly, we construct a hereditary property with an accumulation point at $p = 1/4$. Finally, we derive a significant structural property about those CRGs which occur at accumulation points.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

A *hereditary property* of graphs is a family of graphs which is closed under isomorphism and under vertex-deletion. A *principal hereditary property* is of the form $\text{Forb}(H)$ where H is some fixed graph and $\text{Forb}(H)$ is the family of all graphs which do not contain H as an induced subgraph. In fact, all hereditary properties have the form $\bigcap_{F \in \mathcal{F}} \text{Forb}(F)$ where $\mathcal{F}$ is some family of graphs. In other words, every hereditary property is defined by some collection of forbidden induced subgraphs. It will be convenient to compress notation and write

$$\text{Forb}(\mathcal{F}) \stackrel{\text{def}}{=} \bigcap_{F \in \mathcal{F}} \text{Forb}(F).$$

A hereditary property $\mathcal{H}$ is said to be *non-trivial* if it contains graphs of unbounded order. Equivalently, thanks to Ramsey's theorem, $\mathcal{H}$ is non-trivial if $\mathcal{H} = \text{Forb}(\mathcal{F})$ for some family $\mathcal{F}$ which does not contain both a clique and an anti-clique. Observe that if $\mathcal{H}$ is non-trivial, then, in fact, $\mathcal{H}$ contains graphs of every order.

The edit distance function For two graphs G, H on vertex set V , we define the *normalized edit distance* between G and H to be

* Corresponding author.

E-mail addresses: cocox@iastate.edu (C. Cox), rymartin@iastate.edu (R.R. Martin), dam1@iastate.edu (D. McGinnis).

¹ Supported in part through NSF RTG Grant DMS-1839918.

² Supported in part through Simons Collaboration Grants #353292 and #709641.

³ Supported in part through NSF RTG Grant DMS-1839918.

$$\text{dist}(G, H) \stackrel{\text{def}}{=} \frac{|E(G) \Delta E(H)|}{\binom{|V|}{2}},$$

where Δ denotes the symmetric difference. Informally, $\text{dist}(G, H)$ measures the number of “edits” (that is, edge-additions plus edge-deletions) required to transform G into H .

If $\mathcal{H}$ is a non-trivial hereditary property, then we may extend this definition to discuss the edit distance between G and $\mathcal{H}$ by

$$\text{dist}(G, \mathcal{H}) \stackrel{\text{def}}{=} \min\{\text{dist}(G, H) : H \in \mathcal{H} \text{ s.t. } V(H) = V(G)\}.$$

Note that $\text{dist}(G, \mathcal{H})$ is well-defined since $\mathcal{H}$ is closed under isomorphism and contains graphs of every order. Finally, the *edit distance function* of $\mathcal{H}$ is defined by

$$\text{ed}_{\mathcal{H}}(p) \stackrel{\text{def}}{=} \limsup_{n \rightarrow \infty} \max\left\{\text{dist}(G, \mathcal{H}) : |V(G)| = n, |E(G)| = \left\lfloor p \binom{n}{2} \right\rfloor\right\}, \quad \text{for } p \in [0, 1].$$

Informally, $\text{ed}_{\mathcal{H}}(p)$ records the maximum number of edits required to transform any (large) graph of density p into a member of $\mathcal{H}$.

The edit distance function was formally defined by Balogh–Martin [2], where it was shown that $\text{ed}_{\mathcal{H}}$ is concave down. In a different context, Marchant–Thomason [6] showed that $\text{ed}_{\mathcal{H}}$ is continuous. For a more detailed discussion of the edit distance function and an in-depth review of known results, see [8].

The first result in this paper concerns the behavior of the edit distance function near $p = 0$ and near $p = 1$.

For a family of graphs $\mathcal{F}$, define $\chi(\mathcal{F}) \stackrel{\text{def}}{=} \min\{\chi(F) : F \in \mathcal{F}\}$ and $\overline{\chi}(\mathcal{F}) \stackrel{\text{def}}{=} \min\{\overline{\chi}(F) : F \in \mathcal{F}\}$ where $\overline{\chi}$ denotes the clique-cover number. Suppose that $\mathcal{F}$ does not contain an anti-clique (i.e. $\chi(\mathcal{F}) \geq 2$). By selecting a uniformly random $(\chi(\mathcal{F}) - 1)$ -coloring of the vertices of a graph and deleting any edges induced by each color class, we observe that the resulting graph is $\mathcal{F}$ -free; thus we may bound

$$\text{ed}_{\text{Forb}(\mathcal{F})}(p) \leq \frac{p}{\chi(\mathcal{F}) - 1}.$$

While this bound is not tight for many families $\mathcal{F}$, any known improvements are minuscule when p is sufficiently small. We show that this is no accident:

Theorem 1.1. *Let $\mathcal{F}$ be a family of graphs. If $\mathcal{F}$ does not contain an anti-clique, then*

$$\lim_{p \rightarrow 0^+} \frac{1}{p} \text{ed}_{\text{Forb}(\mathcal{F})}(p) = \frac{1}{\chi(\mathcal{F}) - 1}.$$

Similarly, if $\mathcal{F}$ does not contain a clique, then

$$\lim_{p \rightarrow 1^-} \frac{1}{1 - p} \text{ed}_{\text{Forb}(\mathcal{F})}(p) = \frac{1}{\overline{\chi}(\mathcal{F}) - 1}.$$

In other words, this result determines the slope of the edit distance function at 0 and at 1, provided the forbidden family avoids cliques and anti-cliques.

Remark 1.2. If $\mathcal{F}$ contains an anti-clique (resp. clique), then the slope of $\text{ed}_{\text{Forb}(\mathcal{F})}$ at 0 (resp. 1) is a less interesting question. This is because, if $\mathcal{F}$ contains an anti-clique (resp. clique), then Proposition 2.4 gives that one can compute $\text{ed}_{\text{Forb}(\mathcal{F})}$ over the interval $[0, 1/2]$ (resp. $[1/2, 1]$) via a finite number of structures called p -core CRGs, which we describe below.

Computing the edit distance function A *colored regularity graph* (CRG) K is a clique, together with a partition of its vertices into black and white $V(K) = VB(K) \sqcup VW(K)$ and a partition of its edges into black, white and gray $E(K) = EB(K) \sqcup EW(K) \sqcup EG(K)$.

The term CRG was, to our knowledge, coined by Alon and Stav [1] but the idea traces back throughout studies of hereditary properties using Szemerédi’s regularity lemma [11] such as that by Bollobás and Thomason [3,4].

For a CRG K and a graph F , we say that $F \mapsto K$ if there is a function $\phi : V(F) \rightarrow V(K)$ such that

- For any $uv \in E(F)$, either $\phi(u) = \phi(v) \in VB(K)$ or $\phi(u)\phi(v) \in EB(K) \cup EG(K)$.
- For any $uv \notin E(F)$, either $\phi(u) = \phi(v) \in VW(K)$ or $\phi(u)\phi(v) \in EW(K) \cup EG(K)$.

This notion extends naturally to families of graphs, where we say that $\mathcal{F} \mapsto K$ if there is some $F \in \mathcal{F}$ with $F \mapsto K$. Otherwise, we say that $\mathcal{F} \not\mapsto K$. For a nontrivial hereditary property $\mathcal{H} = \text{Forb}(\mathcal{F})$, denote by $\mathcal{K}(\mathcal{H})$ the set of all CRGs K for which $\mathcal{F} \not\mapsto K$.

For $p \in [0, 1]$, we associate to a CRG K a matrix $M_K(p) \in \mathbb{R}^{V(K) \times V(K)}$ defined by

$$(M_K(p))_{xy} \stackrel{\text{def}}{=} \begin{cases} p & \text{if } x = y \in VW(K) \text{ or if } xy \in EW(K), \\ 1 - p & \text{if } x = y \in VB(K) \text{ or if } xy \in EB(K), \\ 0 & \text{if } xy \in EG(K). \end{cases}$$

From $M_K(p)$, we derive the function

$$g_K(p) \stackrel{\text{def}}{=} \min\{\langle \mu, M_K(p)\mu \rangle : \mu \in \Delta^K\},$$

where $\Delta^K \subseteq [0, 1]^{V(K)}$ is the set of all probability masses on $V(K)$ and $\langle \cdot, \cdot \rangle$ is the standard scalar product. Note that Δ^K is naturally identified with the $(|V(K)| - 1)$ -dimensional simplex.

For any non-trivial hereditary property $\mathcal{H}$ and any $p \in [0, 1]$,

$$\text{ed}_{\mathcal{H}}(p) = \inf_{K \in \mathcal{K}(\mathcal{H})} g_K(p) = \min_{K \in \mathcal{K}(\mathcal{H})} g_K(p). \quad (1)$$

Balogh–Martin [2] proved the first equality and Marchant–Thomason [6] proved the second. In particular, for any fixed $p \in [0, 1]$, the edit distance function of $\mathcal{H}$ is determined by a single CRG.

Accumulation points of the edit distance function Inspired by the known edit distance functions, Martin [8] conjectured the following extension of eq. (1):

Conjecture 1.3 (Martin [8, Conjecture 1]). *For any $\varepsilon > 0$ and any non-trivial hereditary property $\mathcal{H}$, there is some finite subset $\mathcal{K}' \subseteq \mathcal{K}(\mathcal{H})$ such that*

$$\text{ed}_{\mathcal{H}}(p) = \min_{K \in \mathcal{K}'} g_K(p), \quad \text{for all } p \in (\varepsilon, 1 - \varepsilon).$$

In order to approach questions of this form, we introduce the following definition.

Definition 1.4 (Regular points and accumulation points of the edit distance function). Let $\mathcal{H}$ be a non-trivial hereditary property. A point $p_0 \in [0, 1]$ is said to be a *regular point* of $\mathcal{H}$ if there is some $\varepsilon > 0$ and a finite subset $\mathcal{K}' \subseteq \mathcal{K}(\mathcal{H})$ such that

$$\text{ed}_{\mathcal{H}}(p) = \min_{K \in \mathcal{K}'} g_K(p), \quad \text{for every } p \in (p_0 - \varepsilon, p_0 + \varepsilon) \cap [0, 1].$$

Conversely, p_0 is said to be an *accumulation point* of $\mathcal{H}$ if p_0 is not a regular point. Informally, p_0 is an accumulation point of $\mathcal{H}$ if infinitely many CRGs are required to determine $\text{ed}_{\mathcal{H}}$ in any (relatively) open interval containing p_0 .

Appealing to compactness, Conjecture 1.3 is equivalent to the statement: If $\mathcal{H}$ is a non-trivial hereditary property, then every point in $(0, 1)$ is a regular point of $\mathcal{H}$.

Martin and Riasanovsky [10, Theorem 39] proved that $\mathcal{H}$ does not have any accumulation points in the interval $(1 - \varphi^{-1}, \varphi^{-1}) = (2 - \varphi, \varphi - 1) = (0.382 \dots, 0.618 \dots)$ where $\varphi = 1.618 \dots$ is the golden ratio. However, in this paper, we show Conjecture 1.3 to be false:

Theorem 1.5. $p = 1/4$ is an accumulation point of $\text{Forb}(\{K_{1,4}, C_5, C_6, C_7, \dots\})$.

Martin [8, Question 5] additionally asked if $p = 0$ is indeed a accumulation point of $\text{Forb}(K_{3,3})$ and of $\text{Forb}(K_{2,t})$ for $t \geq 9$. We answer this question affirmatively.

Theorem 1.6. $p = 0$ is an accumulation point of $\text{Forb}(K_{3,3})$ and of $\text{Forb}(K_{2,t})$ for all $t \geq 9$.

This result follows from a general classification for when $p = 0$ is an accumulation point of a hereditary property (see Theorem 4.1).

Beyond this, we show that $\text{Forb}(K_{3,3})$ has no other accumulation points.

Theorem 1.7. For any $t \in \mathbb{N}$, $\text{Forb}(K_{t,t})$ has no accumulation points in the interval $(0, 1)$.

In order to prove the above theorem, we establish a significant structural property about those CRGs which occur at accumulation points for a general hereditary property (see Theorem 4.18). We hope that this result can be beneficial to future researchers; however, it requires considerable set-up, so we are unable to state this result in the introduction.

Unfortunately, we were unable to rule out the possibility that $\text{Forb}(K_{2,t})$ has additional accumulation points, though we believe that it does not.

Outline of the paper We begin the paper by recalling background information about the edit distance function and CRGs in Section 2. In Section 3 we establish Theorem 1.1; in order to do so, we prove a variant of the Erdős–Stone theorem. We then prove Theorems 1.5, 1.6 and 1.7 in Section 4, where we additionally establish a significant structural property about those CRGs which occur at accumulation points of a general hereditary property. Finally, we conclude with a few remarks and open questions in Section 5.

2. Preliminaries

In this section, we recall necessary definitions and results about CRGs and the edit distance function.

Proposition 2.1 (Martin [7, Theorem 10v]). *If $\mathcal{H} = \text{Forb}(\mathcal{F})$ is a non-trivial hereditary property, then for any $p \in [0, 1]$,*

$$\text{ed}_{\mathcal{H}}(p) = \text{ed}_{\overline{\mathcal{H}}}(1 - p), \quad \text{where } \overline{\mathcal{H}} = \text{Forb}(\{\overline{F} : F \in \mathcal{F}\}).$$

This, in particular, allows us to restrict our attention to the interval $[0, 1/2]$, since any results in this interval may be immediately translated to a result on $[0, 1]$. This is a common approach in the literature, and we employ it throughout the remainder of this paper.

Definition 2.2. Let K be a CRG and let $p \in (0, 1)$. K is said to be a p -core CRG if

$$g_{K'}(p) > g_K(p)$$

for every proper sub-CRG $K' \subset K$.

Equivalently, K is a p -core CRG if there is a unique $\mu \in \Delta^K$ which achieves $g_K(p)$ and this μ has full support [7, Proposition 12].

Note that any CRG K has a p -core sub-CRG L satisfying $g_L(p) = g_K(p)$.

The above definition makes sense for $p \in \{0, 1\}$ as well, yet it is not useful since the edit distance function is easy to compute for these values. We will therefore define p -core CRGs for $p \in \{0, 1\}$ structurally; these definitions are justified by a limiting argument (see Remark 2.5). Explicitly, Proposition 2.4 shows that the definition of 0-core characterizes p -core CRGs for all $p \in (0, 1/2]$ and that the definition of 1-core characterizes p -core CRGs for all $p \in [1/2, 1)$. Note that Proposition 2.4 also implies that all edges of a $1/2$ -core CRG are gray.

Definition 2.3. Let K be a CRG.

- We say that K is 0-core if K consists only of white and gray edges and $EW(K) \subseteq \binom{VB(K)}{2}$.
- We say that K is 1-core if K consists only of black and gray edges and $EB(K) \subseteq \binom{VW(K)}{2}$.

Proposition 2.4. Fix $p \in (0, 1)$ and let K be a p -core CRG.

1. (Marchant–Thomason [6, Lemma 3.23]) If $p \in (0, 1/2]$, then K is 0-core.
2. (Marchant–Thomason [6, Lemma 3.23]) If $p \in [1/2, 1)$, then K is 1-core.
3. (Martin [7, Lemma 15]) Suppose that $p \in (0, 1/2]$ and that $\mu \in \Delta^K$ achieves $g_K(p)$. For every $u \in VW(K)$, we have

$$\mu(u) = \frac{g_K(p)}{p}.$$

Furthermore, for every $u \in VB(K)$, the weighted-gray-degree of u satisfies

$$d_G(u) \stackrel{\text{def}}{=} \sum_{v \in V(K): uv \in EG(K)} \mu(v) = \frac{p - g_K(p)}{p} + \frac{1 - 2p}{p} \mu(u).$$

We remark that $d_G(u) \leq 1 - \mu(u)$ for any $u \in V(K)$ since $uu \notin EG(K)$ for any vertex u . Additionally, there is an analogue to part 3 in the case that $p \in [1/2, 1)$, though we will not require it thanks to Proposition 2.1.

Remark 2.5. If K is a 0-core CRG, then it can be shown that there is some $p_0 > 0$ such that K is p -core for all $0 < p < p_0$. Similarly, if K is a 1-core CRG, then there is some $p_1 > 0$ such that K is p -core for all $1 - p_1 < p < 1$.

This observation is not used to prove any of the results in this paper, so we leave the proof to the reader. We simply include this remark to further justify the definition of 0-core and 1-core CRGs.

Recall that for a non-trivial hereditary property $\mathcal{H} = \text{Forb}(\mathcal{F})$, we used $\mathcal{K}(\mathcal{H})$ to denote the set of all CRGs K for which $\mathcal{F} \not\vdash K$. We next introduce analogous notation for the CRGs in $\mathcal{K}(\mathcal{H})$ which are also p -core for a given p .

Definition 2.6. Let $\mathcal{H} = \text{Forb}(\mathcal{F})$ be a non-trivial hereditary property. For $p \in [0, 1]$, we denote by $\mathcal{K}_p(\mathcal{H})$ the set of all p -core CRGs K for which $\mathcal{F} \not\vdash K$.

The most important class of CRGs are those consisting only of gray edges.

Definition 2.7. For non-negative integers w, b , the CRG $K(w, b)$ consists of w white vertices and b black vertices and has all gray edges.

Proposition 2.8 (Martin [7, Theorem 7]). Let K and L be CRGs. If $K \oplus L$ denotes the CRG formed by connecting K and L with all gray edges, then

$$\frac{1}{g_{K \oplus L}(p)} = \frac{1}{g_K(p)} + \frac{1}{g_L(p)}.$$

In particular:

1. For $0 < p < 1$,

$$g_{K(w,b)}(p) = \left(\frac{w}{p} + \frac{b}{1-p} \right)^{-1}.$$

2. Let K be any 0-core CRG. If K has w white vertices and K' is the sub-CRG induced by the black vertices of K , then

$$g_K(p) = \left(\frac{w}{p} + \frac{1}{g_{K'}(p)} \right)^{-1}, \quad \text{for all } p \in (0, 1).$$

Intuitively, replacing a non-gray edge of a CRG K with a gray edge can only decrease g_K . This is formally summarized in the following statement.

Claim 2.9. Fix $p \in (0, 1)$, let K be a p -core CRG and suppose that there is some $x \neq y \in V(K)$ with $xy \notin EG(K)$. If K' is the CRG formed by re-coloring the edge xy gray, then $g_K(p) > g_{K'}(p)$.

Proof. Let $\mu \in \Delta^K$ be the probability mass achieving $g_K(p)$; since K is p -core we know that μ has full support. Now, it is easy to see that

$$g_{K'}(p) \leq \langle \mu, M_{K'}(p)\mu \rangle = \langle \mu, M_K(p)\mu \rangle - 2(M_K(p))_{xy}\mu(x)\mu(y) < \langle \mu, M_K(p)\mu \rangle = g_K(p). \quad \square$$

We will explicitly require the following consequence that holds for all CRGs, not just those that are p -core.

Proposition 2.10. Let K be any CRG and fix $p \in (0, 1)$. If $|VW(K)| \leq w$ and $|VB(K)| \leq b$,

$$g_K(p) \geq g_{K(w,b)}(p).$$

Furthermore, the above inequality is strict if one of the following holds:

- $|VW(K)| < w$, or
- $|VB(K)| < b$, or
- K is p -core and has at least one non-gray edge.

Proof. Comparing with part 1 of Proposition 2.8, we observe that $g_{K(w,b)}(p) > g_{K(w',b')}(p)$ whenever either $w < w'$ and $b \leq b'$ or $w \leq w'$ and $b < b'$. Thus the full claim follows by applying Claim 2.9 to any p -core sub-CRG $L \subseteq K$ with $g_L(p) = g_K(p)$. $\square$

3. Slope of the edit distance function at zero

This section is dedicated to proving Theorem 1.1. Appealing to the symmetry established in Proposition 2.1, it suffices to show only that if $\mathcal{F}$ is a family of graphs not containing an anti-clique, then

$$\lim_{p \rightarrow 0^+} \frac{1}{p} \text{ed}_{\text{Forb}(\mathcal{F})}(p) = \frac{1}{\chi(\mathcal{F}) - 1}.$$

Certainly $\mathcal{F} \not\vdash K(\chi(\mathcal{F}) - 1, 0)$, so

$$\text{ed}_{\text{Forb}(\mathcal{F})}(p) \leq g_{K(\chi(\mathcal{F})-1,0)}(p) = \frac{p}{\chi(\mathcal{F}) - 1}, \quad \text{for all } p \in [0, 1];$$

thus the upper bound is clear. Furthermore, observe that $\text{ed}_{\text{Forb}(\mathcal{F})}(p) \geq \text{ed}_{\text{Forb}(F)}(p)$ for any $F \in \mathcal{F}$. Since we can always find an $F \in \mathcal{F}$ for which $\chi(F) = \chi(\mathcal{F})$, in order to prove Theorem 1.1, it suffices to show the following:

Theorem 3.1. *Let H be a graph with $\chi = \chi(H) \geq 2$. For any $\varepsilon > 0$, there is some $p_0 > 0$ such that*

$$\text{ed}_{\text{Forb}(H)}(p) \geq (1 - \varepsilon) \frac{p}{\chi - 1}, \quad \text{whenever } 0 < p < p_0.$$

The main ingredient in the proof of Theorem 3.1 is the following variant of the Erdős–Stone theorem [5]. Note that this result is purely graph-theoretic and does not involve the edit distance function and CRGs.

Lemma 3.2. *Let H be a graph, set $\chi = \chi(H)$ and fix $\varepsilon > 0$. There exists an integer $n_0 = n_0(H, \varepsilon)$ and a number $\alpha = \alpha(H, \varepsilon) > 0$ such that the following holds:*

If G is a graph on $n > n_0$ vertices with

$$\delta(G) > \left(1 - \frac{1}{\chi - 1} + \varepsilon\right)n,$$

and $c: V(G) \rightarrow X$ is any coloring with $|c^{-1}(x)| < \alpha n$ for each $x \in X$, then G contains a (not-necessarily-induced) copy of H , all of whose vertices are colored differently under c .

Proof. Let $K_r(t)$ denote the complete r -partite graph wherein each part has t vertices. Since H is a subgraph of $K_\chi(|V(H)|)$ it suffices to prove the claim only for the graphs $K_r(t)$.

We prove by induction on r that there are appropriate constants $n_0(r, t) \stackrel{\text{def}}{=} n_0(K_r(t), \varepsilon)$ and $\alpha(r, t) \stackrel{\text{def}}{=} \alpha(K_r(t), \varepsilon)$. Employing the convention that $\frac{1}{0} = +\infty$, we may consider the case of $r = 1$ as our base case. Here, $K_1(t)$ is simply an anti-clique of size t , and so the result follows from setting $\alpha(1, t) = 1/t$ and $n_0(1, t) = t$.

For the induction step, fix $r \geq 2$ and set

$$T \stackrel{\text{def}}{=} \frac{2t}{\varepsilon}, \quad n_0(r, t) \stackrel{\text{def}}{=} n_0(r-1, T), \quad \text{and} \quad \alpha(r, t) \stackrel{\text{def}}{=} \min \left\{ \alpha(r-1, T), \frac{\varepsilon}{2rt} \binom{T}{t}^{-(r-1)t} \right\}. \quad (2)$$

Note that by making ε slightly smaller if necessary, we may suppose that T is an integer; hence the above quantities are well-defined.

Let $G = (V, E)$ be a graph on $n > n_0(r, t)$ vertices with $\delta(G) > (1 - \frac{1}{r-1} + \varepsilon)n$, and let $c: V \rightarrow X$ be any coloring with $|c^{-1}(x)| < \alpha n$ for all $x \in X$, where $\alpha = \alpha(r, t)$. We must show that G contains a copy of $K_r(t)$, all of whose vertices are colored differently under c . From eq. (2) and the induction hypothesis, we know that G contains a copy of $K_{r-1}(T)$, all of whose vertices receive different colors under c . Label the parts of this $K_{r-1}(T)$ as $U_1, \dots, U_{r-1}$.

Define the set

$$A \stackrel{\text{def}}{=} \left\{ (v, W_1, \dots, W_{r-1}) \in V \times \binom{U_1}{t} \times \dots \times \binom{U_{r-1}}{t} : N(v) \supseteq \bigcup_{i=1}^{r-1} W_i \right\},$$

where $N(v)$ denotes the neighborhood of the vertex v . We claim that we have succeeded in finding our desired copy of $K_r(t)$ if $|A| \geq \binom{T}{t}^{r-1} \cdot rt \cdot \alpha n$. Indeed, if this is the case, then there is some $(W_1, \dots, W_{r-1})$ such that $B \stackrel{\text{def}}{=} \{v \in V : (v, W_1, \dots, W_{r-1}) \in A\}$ satisfies $|B| \geq rt \cdot \alpha n$. Since each color class of c has size at most αn , this implies that there is some $B' \subseteq B$ with $|B'| \geq rt$, all of whose vertices have distinct colors under c . Removing vertices of B' that share a color with a vertex in $W_1 \cup \dots \cup W_{r-1}$ leaves at least t vertices; these t vertices along with $W_1, \dots, W_{r-1}$ form our desired copy of $K_r(t)$.

Therefore, the remainder of the proof is dedicated to proving that $|A| \geq \binom{T}{t}^{r-1} \cdot rt \cdot \alpha n$.

For $v \in V$, set $f(v) \stackrel{\text{def}}{=} |N(v) \cap \bigcup_{i=1}^{r-1} U_i|$. We begin by bounding

$$\sum_{v \in V} f(v) = \sum_{u \in \bigcup_{i=1}^{r-1} U_i} \deg(u) > T \cdot (r-1) \cdot \left(1 - \frac{1}{r-1} + \varepsilon\right)n. \quad (3)$$

Now, let S be the collection of all $v \in V$ for which $f(v) \geq T \cdot (r-1) \cdot \left(1 - \frac{1}{r-1} + \frac{\varepsilon}{2}\right)$. Observe that

$$\begin{aligned} \sum_{v \in V} f(v) &\leq |S| \cdot T \cdot (r-1) + (n - |S|) \cdot T \cdot (r-1) \cdot \left(1 - \frac{1}{r-1} + \frac{\varepsilon}{2}\right) \\ &= T \cdot (r-1) \cdot \left[\left(1 - \frac{1}{r-1} + \frac{\varepsilon}{2}\right)n + \left(\frac{1}{r-1} - \frac{\varepsilon}{2}\right)|S| \right]. \end{aligned}$$

Comparing with eq. (3), we thus bound

$$|S| > \frac{r-1}{2} \cdot \varepsilon n \geq \frac{\varepsilon n}{2}.$$

Now, for each $v \in S$, it must be the case that

$$|N(v) \cap U_i| \geq \frac{r-1}{2} \cdot \varepsilon T \geq \frac{\varepsilon T}{2}, \quad \text{for each } i \in [r-1],$$

since otherwise

$$f(v) < T \cdot (r-2) + \frac{r-1}{2} \cdot \varepsilon T = T \cdot (r-1) \cdot \left(1 - \frac{1}{r-1} + \frac{\varepsilon}{2}\right);$$

contradicting the definition of S .

Since $T = 2t/\varepsilon$, we thus have $|N(v) \cap U_i| \geq t$ for all $v \in S$ and all $i \in [r-1]$; therefore,

$$|A| \geq |S| \geq \frac{\varepsilon n}{2}.$$

Combining this inequality with eq. (2), we finally bound

$$\binom{T}{t}^{r-1} \cdot rt \cdot \alpha n \leq \frac{\varepsilon n}{2} \leq |A|,$$

which concludes the proof. $\square$

Returning to the edit distance function and CRGs, we next derive a simple consequence of Proposition 2.4.

Claim 3.3. Fix p, ε such that $1/2 > \varepsilon > p > 0$ and $r \in \mathbb{N}$ and let K be a p -core CRG. Suppose that $g_K(p) < (1 - \varepsilon)\frac{p}{r}$ and let $\mu \in \Delta^K$ be the probability mass (with full support) achieving $g_K(p)$. Then for each $v \in V(K)$,

$$d_G(v) > 1 - \frac{1}{r} + \frac{\varepsilon}{r}, \quad \text{and} \quad \mu(v) < \frac{p}{r}.$$

Proof. Set $g = g_K(p)$. From part 3 of Proposition 2.4, we know that

$$d_G(v) = \frac{p-g}{p} + \frac{1-2p}{p} \mu(v) > \frac{p-g}{p} > 1 - \frac{1-\varepsilon}{r},$$

which establishes the first piece of the claim. For the second part of the claim, since $d_G(v) \leq 1 - \mu(v)$ and $p < \varepsilon$, we obtain

$$\frac{p-g}{p} + \frac{1-2p}{p} \mu(v) \leq 1 - \mu(v) \implies \mu(v) \leq \frac{g}{1-p} < \frac{1-\varepsilon}{1-p} \cdot \frac{p}{r} < \frac{p}{r}. \quad \square$$

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. Suppose the claim is false and let H be a counterexample with $\chi = \chi(H) \geq 2$ minimum. Since H is a counterexample, there is some $\varepsilon > 0$ and some monotone sequence $p_n \rightarrow 0$ such that, for each $n \in \mathbb{N}$, there is a p_n -core CRG Γ_n for which

- I. $H \not\rightarrow \Gamma_n$, and

$$\text{II. } \text{ed}_{\text{Forb}(H)}(p_n) = g_{\Gamma_n}(p_n) < (1 - \varepsilon) \frac{p_n}{\chi - 1}.$$

By passing to a subsequence if necessary, we can ensure the following:

Claim 3.4. *Each Γ_n has only black vertices.*

Proof. First, we know that $|VW(\Gamma_n)| < \chi(H)$ for each n , so by passing to a subsequence if necessary, we may suppose that $|VW(\Gamma_n)| = w$ for each $n \in \mathbb{N}$. Suppose for the sake of contradiction that $w > 0$ and let Γ'_n be the sub-CRG induced by the black vertices of Γ_n . Note that Γ'_n is nonempty since otherwise $\Gamma_n = K(w, 0)$ which has $g_{\Gamma_n}(p_n) = \frac{p_n}{w} \geq \frac{p_n}{\chi - 1}$; contradicting property II. Applying part 2 of Proposition 2.8, we bound

$$\frac{1}{g_{\Gamma_n}(p_n)} = \frac{w}{p_n} + \frac{1}{g_{\Gamma'_n}(p_n)} \implies g_{\Gamma'_n}(p_n) < (1 - \varepsilon) \frac{p_n}{\chi - w - 1 + \varepsilon w}.$$

Now, if $w = \chi - 1$, then $g_{\Gamma'_n}(p_n) < p_n/\varepsilon$. On the other hand, certainly $|\Gamma'_n| < |V(H)|$ or else $H \mapsto \Gamma_n$, so, through Proposition 2.10,

$$g_{\Gamma'_n}(p_n) \geq g_{K(0, |V(H)|)}(p_n) = \frac{1 - p_n}{|V(H)|} > \frac{p_n}{\varepsilon}$$

for n sufficiently large; a contradiction.

We must therefore have $1 \leq w \leq \chi - 2$. Let H' be any induced subgraph of H which has $\chi(H') = \chi - w$; we know that $2 \leq \chi(H') \leq \chi - 1$. Now, it must be the case that $H' \not\mapsto \Gamma'_n$ for every n , or else $H \mapsto \Gamma_n$. Thus, for every n ,

$$\text{ed}_{\text{Forb}(H')}(p_n) \leq g_{\Gamma'_n}(p_n) < (1 - \varepsilon) \frac{p_n}{\chi - w - 1 + \varepsilon w} < (1 - \varepsilon) \frac{p_n}{\chi(H') - 1};$$

contradicting the minimality of the graph H . $\square$

We now know that each Γ_n consists only of black vertices. Let μ_n denote the probability mass achieving $g_{\Gamma_n}(p_n)$. Fix n and N_0 sufficiently large to be chosen later; we will consider n to be large, but fixed, whereas we will consider $N_0 \rightarrow \infty$. Define the graph G , which has vertex partition $V(G) = \bigsqcup_{x \in V(\Gamma_n)} V_x$ where $|V_x| = \lfloor \mu_n(x) \cdot N_0 \rfloor$, and whose edge-set consists of all edges between V_x and V_y whenever $xy \in EG(\Gamma_n)$. In other words, G is formed by retaining only the gray edges of Γ_n and then blowing up each vertex $x \in V(\Gamma_n)$ into an independent set of size $\lfloor \mu_n(x) \cdot N_0 \rfloor$. Also, let $c: V(G) \rightarrow V(\Gamma_n)$ be the coloring defined by $c(v) = x$ if $v \in V_x$. We seek to show that G has large minimum degree and that each color class of c is sufficiently small so that we may apply Lemma 3.2.

We begin with a few observations about the structure of G .

Firstly, setting $N \stackrel{\text{def}}{=} |V(G)|$, we certainly have

$$N = \sum_{x \in V(\Gamma_n)} \lfloor \mu_n(x) \cdot N_0 \rfloor \implies N_0 - |V(\Gamma_n)| < N \leq N_0.$$

Next, consider any $v \in V(G)$ and suppose that $c(v) = x$. By applying Claim 3.3 and property II, we bound

$$\begin{aligned} \deg(v) &= \sum_{y \in V(\Gamma_n): xy \in EG(\Gamma_n)} \lfloor \mu_n(x) \cdot N_0 \rfloor \geq d_G(x) \cdot N - |V(\Gamma_n)| \\ &\geq \left(1 - \frac{1}{\chi - 1} + \frac{\varepsilon}{\chi - 1}\right) N - |V(\Gamma_n)|. \end{aligned}$$

Finally, by selecting n sufficiently large so that $p_n < \varepsilon$, for any $x \in V(\Gamma_n)$

$$|c^{-1}(x)| = \lfloor \mu_n(x) \cdot N_0 \rfloor \leq \mu_n(v) \cdot N + |V(\Gamma_n)| < \frac{p_n}{\chi - 1} \cdot N + |V(\Gamma_n)|,$$

thanks again to Claim 3.3 and property II.

Thus, by first selecting n sufficiently large (so that p_n is sufficiently small) and then selecting N_0 (and hence N) sufficiently large compared to $|V(\Gamma_n)|$, we may apply Lemma 3.2 to the graph G to find a copy of H , all of whose vertices are colored differently under c .

Suppose that $\phi: V(H) \rightarrow V(G)$ realizes H as such a subgraph of G and consider the map $c \circ \phi: V(H) \rightarrow V(\Gamma_n)$. Since the vertices of this copy of H are all colored differently under c , we know that $c \circ \phi$ is an injection. We therefore find that $c \circ \phi$ maps edges of H to gray edges of Γ_n and maps non-edges of H to either white or gray edges of Γ_n . This, however, means that $H \mapsto \Gamma_n$; contradicting property I. $\square$

4. Accumulation points

We begin by establishing a necessary and sufficient condition for 0 to be an accumulation point of a hereditary property. Thanks to Proposition 2.1, there is an analogous classification for when 1 is an accumulation point.

Theorem 4.1. *Let $\mathcal{H} = \text{Forb}(\mathcal{F})$ be a non-trivial hereditary property and set $\chi = \chi(\mathcal{F})$. For $p \in (0, 1/2)$, define*

$$q_{\mathcal{H}}(p) \stackrel{\text{def}}{=} \min\{g_K(p) : |VW(K)| = \chi - 1, K \in \mathcal{K}_0(\mathcal{H})\}.$$

$p = 0$ is a regular point of $\mathcal{H}$ if and only if

$$\text{ed}_{\mathcal{H}}(p) = q_{\mathcal{H}}(p), \quad \text{for all } p \text{ sufficiently small.}$$

Proof. Recalling that $\mathcal{K}_0(\mathcal{H})$ is the set of 0-core CRGs in $\mathcal{K}(\mathcal{H})$, we observe that $\{K \in \mathcal{K}_0(\mathcal{H}) : |VW(K)| = \chi - 1\}$ is finite. Indeed, if $F \in \mathcal{F}$ has $\chi(F) = \chi$, then $F \mapsto K$ whenever K is a 0-core CRG, $|VW(K)| \geq \chi - 1$ and $|VB(K)| \geq |V(F)|$. Therefore $q_{\mathcal{H}}(p)$ is well-defined and also the “if” direction is immediate.

To establish the “only if” direction, suppose that $p = 0$ is a regular point of $\mathcal{H}$; thus, we can find some finite family $\mathcal{K}' \subseteq \mathcal{K}(\mathcal{H})$ with the property that

$$\text{ed}_{\mathcal{H}}(p) = \min_{K \in \mathcal{K}'} g_K(p)$$

for all p sufficiently small. Certainly we may assume that $\mathcal{K}' \subseteq \mathcal{K}_0(\mathcal{H})$.

Set

$$N \stackrel{\text{def}}{=} \max_{K \in \mathcal{K}'} |VB(K)|,$$

and consider some $K \in \mathcal{K}'$. If $|VW(K)| = \chi - 1$, then $g_K(p) \geq q_{\mathcal{H}}(p)$ by definition, so suppose that $|VW(K)| = w \leq \chi - 2$. In this case, we apply Proposition 2.10 to bound

$$g_K(p) \geq g_{K(w,N)}(p) = \frac{(1-p)p}{(1-p)w + pN} > \frac{p}{\chi - 1} = g_{K(\chi-1,0)}(p),$$

provided $p < 1/(N+1)$. Since certainly $K(\chi-1,0) \in \mathcal{K}_0(\mathcal{H})$, we have shown that, for all p sufficiently small,

$$q_{\mathcal{H}}(p) \geq \text{ed}_{\mathcal{H}}(p) = \min_{K \in \mathcal{K}'} g_K(p) \geq q_{\mathcal{H}}(p). \quad \square$$

Using Theorem 4.1, we can quickly verify that $p = 0$ is indeed an accumulation point for some hereditary properties.

Proof of Theorem 1.6. First, set $\mathcal{H} = \text{Forb}(K_{3,3})$. If $K \in \mathcal{K}_0(\mathcal{H})$ has exactly one white vertex, then K has at most two black vertices. Since $K_{3,3} \not\rightarrow K(1,2)$, an appeal to Proposition 2.10 implies that

$$q_{\mathcal{H}}(p) = \frac{p(1-p)}{1+p}, \quad \text{for every } p \in (0, 1/2).$$

However, Marchant and Thomason [6, Example 5.19] showed that there is a monotone sequence $p_n \rightarrow 0$ such that

$$\text{ed}_{\mathcal{H}}(p_n) < \frac{p_n(1-p_n)}{1+p_n}, \quad \text{for every } n \in \mathbb{N}.$$

Thus, thanks to Theorem 4.1, we know that $p = 0$ is indeed an accumulation point of $\mathcal{H}$.

Next set $\mathcal{H} = \text{Forb}(K_{2,t})$ for any $t \geq 9$. Observe that if $K \in \mathcal{K}_0(\mathcal{H})$ has exactly one white vertex, then K has at most one black vertex. Since $K_{2,t} \not\rightarrow K(1,1)$ for $t \geq 2$, another appeal to Proposition 2.10 implies that

$$q_{\mathcal{H}}(p) = p(1-p), \quad \text{for every } p \in (0, 1/2).$$

However, Martin and McKay [9, Theorem 8 & Corollary 9] showed that there is a monotone sequence $p_n \rightarrow 0$ such that

$$\text{ed}_{\mathcal{H}}(p_n) < p_n(1-p_n), \quad \text{for every } n \in \mathbb{N}.$$

Thus, again thanks to Theorem 4.1, we know that $p = 0$ is indeed an accumulation point of $\mathcal{H}$. $\square$

We prove that $p = 0$ is the *only* accumulation point of $\text{Forb}(K_{3,3})$ in the proof of Theorem 1.7 in Section 4.1.

Before we get to this, we prove Theorem 1.5. To do so, we need an intermediate result. Firstly, for two CRGs K, L , denote by $K \vee L$ the CRG formed by connecting K and L with all-white edges. Next, denote by $\mathcal{P}_n$ the n -vertex CRG with only black vertices, whose gray edges form an n -vertex path and all of whose other edges are white.

Lemma 4.2. For any $n_1, \dots, n_\ell \in \mathbb{N}$, there is some $\varepsilon > 0$ such that

$$g_{\mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}}(p) > p, \quad \text{for all } p \in (1/4 - \varepsilon, 1/4 + \varepsilon)$$

Proof. Since g_K is a continuous function for any CRG K , it suffices to show only that

$$g_{\mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}}(1/4) > 1/4.$$

Since any sub-CRG of $\mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}$ also has the form $\mathcal{P}_{m_1} \vee \dots \vee \mathcal{P}_{m_t}$ for some $m_1, \dots, m_t \in \mathbb{N}$, we may assume, without loss of generality, that $K \stackrel{\text{def}}{=} \mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}$ is 1/4-core.

For a vertex $v \in V(K)$, denote by $NG(v) \stackrel{\text{def}}{=} \{u \in V(K) : uv \in EG(K)\}$ the gray-neighborhood of v . Set $g = g_K(1/4)$ and let $\mu \in \Delta^K$ be the probability mass achieving g .

By applying part 3 of Proposition 2.4, we know that

$$2\mu(v) + 1 - 4g = \sum_{u \in NG(v)} \mu(u), \quad \text{for all } v \in V(K).$$

Supposing that $|V(K)| = n$, summing both sides over all $v \in V(K)$ then yields

$$\begin{aligned} 2 + n(1 - 4g) &= \sum_{v \in V(K)} \sum_{u \in NG(v)} \mu(u) = \sum_{v \in V(K)} |NG(v)| \cdot \mu(v) \\ &= 2 - \sum_{v: |NG(v)|=1} \mu(v) - 2 \sum_{v: |NG(v)|=0} \mu(v) \\ \implies g &= \frac{1}{4} + \frac{1}{4n} \sum_{v: |NG(v)|=1} \mu(v) + \frac{1}{2n} \sum_{v: |NG(v)|=0} \mu(v) > \frac{1}{4}, \end{aligned}$$

since μ has full support and K certainly has at least one vertex $v \in V(K)$ with $|NG(v)| \in \{0, 1\}$. $\square$

We now demonstrate a hereditary property which has an accumulation point at $p = 1/4$.

Proof of Theorem 1.5. Set $\mathcal{H} = \text{Forb}(\mathcal{F})$ where $\mathcal{F} = \{K_{1,4}, C_5, C_6, \dots\}$. We begin by determining $\mathcal{K}_0(\mathcal{H})$.

Claim 4.3.

$$\mathcal{K}_0(\mathcal{H}) = \{K(1, 0)\} \cup \{\mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell} : n_1, \dots, n_\ell \in \mathbb{N}\}.$$

Proof. Since $\mathcal{F}$ does not contain an anti-clique, we know that $K(1, 0) \in \mathcal{K}_0(\mathcal{H})$. Furthermore, since $K_{1,4} \mapsto K(2, 0)$ and $K_{1,4} \mapsto K(1, 1)$, we observe that $K(1, 0)$ is the only member of $\mathcal{K}_0(\mathcal{H})$ which has at least one white vertex.

Next it is easy to observe that $K_{1,4} \not\mapsto \mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}$ and that if $C_m \mapsto \mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}$, then $m \leq 4$. Thus, all that is left to show is that if $K \in \mathcal{K}_0(\mathcal{H})$ has only black vertices, then $K = \mathcal{P}_{n_1} \vee \dots \vee \mathcal{P}_{n_\ell}$ for some $n_1, \dots, n_\ell \in \mathbb{N}$.

Again, since $K_{1,4} \not\mapsto K$, we observe that every vertex of K is incident to at most two gray edges. Beyond this, if K contains an n -vertex gray-edge cycle, then $C_m \mapsto K$ for every $n \leq m \leq 2n$. We conclude that the gray-edges of K form a linear forest and so we have established the claim. $\square$

We are now ready to show that $p = 1/4$ is an accumulation point of $\mathcal{H}$. Suppose to the contrary that there exists some $\varepsilon > 0$ and some finite family $\mathcal{K}' \subseteq \mathcal{K}(\mathcal{H})$ such that

$$\text{ed}_{\mathcal{H}}(p) = \min_{K \in \mathcal{K}'} g_K(p), \quad \text{for all } p \in [1/4, 1/4 + \varepsilon).$$

Certainly we may suppose that $\varepsilon < 1/4$ and that $\mathcal{K}' \subseteq \mathcal{K}_0(\mathcal{H})$.

Now, we know that $g_{K(1,0)}(p) = p$ for all $p \in [0, 1]$, so by additionally applying Lemma 4.2 and using the fact that $\mathcal{K}'$ is finite, we observe that there is some $\varepsilon_1 > 0$ for which

$$\text{ed}_{\mathcal{H}}(p) = \min_{K \in \mathcal{K}'} g_K(p) \geq p, \quad \text{for all } p \in [1/4, 1/4 + \varepsilon_1). \quad (4)$$

However, for $n \geq 3$, letting $\mathbf{u} \in \Delta^{\mathcal{P}_n}$ denote the uniform distribution, we compute

$$g_{\mathcal{P}_n}(p) \leq \langle \mathbf{u}, M_{\mathcal{P}_n}(1/4)\mathbf{u} \rangle = \frac{1-p}{n} + p \cdot \frac{(n-2)(n-3) + 2(n-1)}{n^2} = p - \frac{4p-1}{n} + \frac{4p}{n^2}.$$

Therefore, since $\mathcal{F} \not\rightarrow \mathcal{P}_n$, we see that for any $p > 1/4$ and n sufficiently large, we have

$$\text{ed}_{\mathcal{H}}(p) \leq g_{\mathcal{P}_n}(p) < p,$$

contradicting eq. (4). $\square$

Remark 4.4. The above proof actually establishes a slightly stronger claim; namely, if $\mathcal{C}$ is any collection of cycles satisfying

- If $C_m \in \mathcal{C}$, then $m \geq 5$, and
- For every $n \geq 3$, there is some $n \leq m \leq 2n$ for which $C_m \in \mathcal{C}$,

then $\text{Forb}(\{K_{1,4}\} \cup \mathcal{C})$ has an accumulation point at $p = 1/4$.

4.1. CRGs at accumulation points

In this section, we establish structural properties of CRGs at accumulation points of the edit distance function.

In order to do so, we will require the framework established by Marchant and Thomason [6]. In their paper, Marchant and Thomason work with a function different from the edit distance function, which they show to be equivalent to $1 - \text{ed}_{\mathcal{H}}$ (see [6, Section 2.2]). Moreover, they use somewhat different notation. Below we translate their framework to our situation of edit distance, and we modify their notation to be more consistent with our own.

We first define a class of objects closely related to CRGs.

Definition 4.5. A *colored graph* G is a clique with a coloring of its edges using the colors black, white and gray: $E(G) = EB(G) \sqcup EW(G) \sqcup EG(G)$. We use $|G|$ to denote the number of vertices of a colored graph G .

For a colored graph G and a number $p \in [0, 1]$, define the following quantities:

- For an edge $e \in E(G)$, the p -weight of e is defined as

$$w_p(e) \stackrel{\text{def}}{=} \begin{cases} p & \text{if } e \in EW(G), \\ 1 - p & \text{if } e \in EB(G), \\ 0 & \text{if } e \in EG(G). \end{cases}$$

- For $v \in V(G)$, the p -degree of v is defined as

$$d_p(v) \stackrel{\text{def}}{=} \sum_{u \in V(G)} w_p(uv).$$

- Finally, the maximum- p -degree of G is defined as

$$\Delta_p(G) \stackrel{\text{def}}{=} \max_{v \in V(G)} d_p(v).$$

The only difference between a colored graph and a CRG is that a colored graph does not come equipped with a coloring of its vertices.

Definition 4.6. For two colored graphs G, H , we say that $G \sqsubseteq H$ if there is an injection $\phi: V(G) \rightarrow V(H)$ so that

- If $uv \in EB(G)$, then $\phi(u)\phi(v) \in EB(H) \cup EG(H)$, and
- If $uv \in EW(G)$, then $\phi(u)\phi(v) \in EW(H) \cup EG(H)$, and
- If $uv \in EG(G)$, then $\phi(u)\phi(v) \in EG(H)$.

In other words, $G \sqsubseteq H$ if G is formed from H by deleting some vertices and recoloring some gray edges with either white or black. Note that $\sqsubseteq$ is a partial-order on the set of colored graphs and that $G \sqsubseteq H \implies G \sqsubseteq H$.

A single colored graph is uninteresting on its own – we instead need to work with sequences of colored graphs.

Definition 4.7. A *colored graph sequence* is a sequence of colored graphs $(G_n)_{n \in \mathbb{N}}$ of unbounded order.

It is useful to think of a colored graph sequence as being an approximation of a CRG, though we will see later that colored graph sequences can encode more intricate information.

From a single CRG, we can generate a colored graph sequence by “blowing-up” the vertices into colored cliques:

Definition 4.8. Let K be a CRG and let $m \in \mathbb{N}$. The colored graph $m \times K$ has vertex-set $V = \bigsqcup_{x \in V(K)} V_x$ where $|V_x| = m$ for every $x \in V(K)$. Furthermore, all edges between vertices in V_x have the same color as the vertex $x \in V(K)$, and all edges between V_x and V_y have the same color as the edge $xy \in E(K)$.

In the opposite direction, the following definition lays out a template for passing from a colored graph sequence to a CRG.

Definition 4.9. Let $(G_n)_{n \in \mathbb{N}}$ be a colored graph sequence and let K be a CRG. We say that $K \sqsubset (G_n)$ if for every $m \in \mathbb{N}$, there is some $n_m \in \mathbb{N}$ for which $m \times K \subseteq G_{n_m}$.

Observe that if $K \sqsubset (G_n)$, then for any $m \in \mathbb{N}$, there are, in fact, infinitely many n 's for which $m \times K \subseteq G_n$. Furthermore, certainly if $K \subseteq L$ and $L \sqsubset (G_n)$, then also $K \sqsubset (G_n)$.

To motivate the above definition, consider a graph H ; we may identify H with a colored graph by coloring the edges of H black and coloring the non-edges of H white. Observe that for a CRG K , if $H \mapsto K$, then $H \subseteq m \times K$ for m sufficiently large (namely, $m \geq |V(H)|$). Therefore, if (G_n) is a colored graph sequence with the property that $H \not\subseteq G_n$ for all n , then for any CRG K with $K \sqsubset (G_n)$, it must be the case that $H \not\mapsto K$.

We will require the following results of Marchant and Thomason from [6]. First, a CRG L is said to be an *extension* of a CRG K if K is obtained by deleting a single vertex from L .

Lemma 4.10 (Marchant–Thomason [6, Lemma 3.11]). Let $(G_n)_{n \in \mathbb{N}}$ be a colored graph sequence and suppose that K is a CRG for which $K \sqsubset (G_n)$. Then for any $\mu \in \Delta^K$, there is an extension L of K for which $L \sqsubset (G_n)$ and the vertex $\{v\} = V(L) \setminus V(K)$ satisfies

$$\sum_{u \in V(K)} \mu(u) w_p(uv) \leq \limsup_{n \rightarrow \infty} \frac{\Delta_p(G_n)}{|G_n|}.$$

Lemma 4.11 (Marchant–Thomason [6, Lemmas 3.13 & 3.22]). Fix $p \in (0, 1/2)$, let K be a p -core CRG and let $\mu \in \Delta^K$ be the probability mass achieving $g_K(p)$. Suppose that L is an extension of K such that

$$\sum_{u \in V(K)} \mu(u) w_p(uv) < g_K(p),$$

where $\{v\} = V(L) \setminus V(K)$. If L' is any sub-CRG of L which is p -core and has $g_{L'}(p) = g_L(p)$, then $g_{L'}(p) < g_K(p)$ and $|VW(L')| \geq |VW(K)|$.

The bulk of the arguments in this section are built on variants of the following observation, which, informally, states that a large number of black vertices of a 0-core CRG behave as if they were a single white vertex.

Proposition 4.12. Let K be a 0-core CRG and fix $n \in \mathbb{N}$. If $|VW(K)| = t$ and $|VB(K)| \geq n$, then $n \times K(t+1, 0) \subseteq n \times K$.

Proof. Let $V_1, \dots, V_{t+1}$ be the vertex partition of $V(n \times K(t+1, 0))$ defined in Definition 4.8; label $V_i = \{v_i^1, \dots, v_i^n\}$. Additionally, let $W_1, \dots, W_t, B_1, \dots, B_m$ be the vertex partition of $V(n \times K)$ where $W_1, \dots, W_t$ correspond to the white vertices of K and $B_1, \dots, B_m$ correspond to the black vertices of K . Label $W_i = \{w_i^1, \dots, w_i^n\}$ and $B_i = \{b_i^1, \dots, b_i^n\}$.

Define the map $\phi: V(n \times K(t+1, 0)) \rightarrow V(n \times K)$ by

$$\phi(v_i^j) = \begin{cases} w_i^j & \text{if } i \in [t], \\ b_j^1 & \text{if } i = t+1, \end{cases}$$

which is well-defined and an injection since $m \geq n$. Since K is 0-core, all edges induced by $\phi(V_i)$ for any $i \in [t+1]$ are either white or gray, and all edges between $\phi(V_i)$ and $\phi(V_j)$ for $i \neq j \in [t+1]$ are gray. Therefore $n \times K(t+1, 0) \subseteq n \times K$. $\square$

The following result was implicitly established by Marchant and Thomason in the proof of Theorem 3.25 in [6].

Lemma 4.13. Fix $p \in (0, 1/2)$, let $(G_n)_{n \in \mathbb{N}}$ be a colored graph sequence and set

$$w = \sup\{t \in \mathbb{N} \cup \{0\} : K(t, 0) \sqsubset (G_n)\}.$$

If w is finite, then there is a p -core CRG K such that $K \sqsubset (G_n)$, K has at least w white vertices and

$$g_K(p) \leq \limsup_{n \rightarrow \infty} \frac{\Delta_p(G_n)}{|G_n|}.$$

Proof. Set

$$\Delta_p \stackrel{\text{def}}{=} \limsup_{n \rightarrow \infty} \frac{\Delta_p(G_n)}{|G_n|}.$$

Suppose for the sake of contradiction that the claim is false; we aim to show that $K(w+1, 0) \sqsubset (G_n)$. Under this assumption, we begin by defining a sequence of p -core CRGs $(\Gamma_n)_{n \in \mathbb{N}}$ such that

- $|VW(\Gamma_n)| \geq w$, and
- $\Gamma_n \sqsubset (G_n)$, and
- $g_{\Gamma_n}(p) > g_{\Gamma_{n+1}}(p)$.

Begin with $\Gamma_1 = K(w, 0)$, which is p -core, has w white vertices and $\Gamma_1 \sqsubset (G_n)$ by definition. Assuming Γ_n has been defined, we define Γ_{n+1} as follows:

Let L be an extension of Γ_n guaranteed by Lemma 4.10, so $L \sqsubset (G_n)$ and

$$\sum_{u \in V(\Gamma_n)} \mu(u) w_p(uv) \leq \Delta_p < g_{\Gamma_n}(p),$$

where $\{v\} = V(L) \setminus V(\Gamma_n)$ and the latter inequality holds by assumption. Now, let Γ_{n+1} be any p -core sub-CRG of L with $g_{\Gamma_{n+1}}(p) = g_L(p)$. Since $\Gamma_{n+1} \subseteq L$ and $L \sqsubset (G_n)$, we must have $\Gamma_{n+1} \sqsubset (G_n)$. Additionally, thanks to Lemma 4.11, we know that $g_{\Gamma_{n+1}}(p) < g_{\Gamma_n}(p)$ and that $|VW(\Gamma_{n+1})| \geq |VW(\Gamma_n)| \geq w$. Thus, Γ_{n+1} satisfies the claimed properties.

We claim now that $K(w+1, 0) \sqsubset (G_n)$, which will contradict the definition of w and conclude the proof. Firstly, if $|VW(\Gamma_n)| > w$ for some $n \in \mathbb{N}$, then $K(w+1, 0) \subseteq \Gamma_n$ and so $K(w+1, 0) \sqsubset (G_n)$. We may thus suppose that $|VW(\Gamma_n)| = w$ for each $n \in \mathbb{N}$.

Fix any $N \in \mathbb{N}$. Since $g_{\Gamma_1}(p) > g_{\Gamma_2}(p) > \dots$, we know that the Γ_n 's are distinct; thus, since each Γ_n has exactly w white vertices, there must be some $t \in \mathbb{N}$ for which $|VB(\Gamma_t)| \geq N$. Now, since $\Gamma_t \sqsubset (G_n)$, there is some $t' \in \mathbb{N}$ for which $N \times \Gamma_t \subseteq G_{t'}$. Thus, by applying Proposition 4.12, we have $N \times K(w+1, 0) \subseteq N \times \Gamma_t \subseteq G_{t'}$. Since $N \in \mathbb{N}$ was arbitrary, we conclude that $K(w+1, 0) \sqsubset (G_n)$, which concludes the proof. $\square$

We next demonstrate a more refined construction of a colored graph sequence from a single CRG wherein we “blow-up” the vertices into colored cliques of various sizes.

Definition 4.14. Let K be a CRG and fix any $\mu \in \Delta^K$. For a positive integer n , denote by $K[\mu, n]$ the colored graph with vertex-partition $V(K[\mu, n]) = \bigsqcup_{x \in V(K)} V_x$ where $|V_x| = \lfloor \mu(x) \cdot n \rfloor$ for every $x \in V(K)$. Furthermore, all edges between vertices in V_x have the same color as the vertex $x \in V(K)$, and all edges between V_x and V_y have the same color as the edge $xy \in E(K)$.

Notice that if $k = |V(K)|$, then $m \times K = K[\mathbf{u}, km]$ where $\mathbf{u} \in \Delta^K$ is the uniform distribution.

Sequences of colored graphs constructed in this manner mirror the properties of the original CRG.

Proposition 4.15. Fix $p \in (0, 1/2]$. If K is a p -core CRG and $\mu \in \Delta^K$ is the probability mass achieving $g_K(p)$, then

$$g_K(p) = \lim_{n \rightarrow \infty} \frac{\Delta_p(K[\mu, n])}{|K[\mu, n]|}.$$

Proof. For notational ease, set $G_n \stackrel{\text{def}}{=} K[\mu, n]$, and note that

$$|G_n| = \sum_{x \in V(K)} \lfloor \mu(x) \cdot n \rfloor \implies |G_n| = n + O(1),$$

for n sufficiently large.

Fix any $v \in V(G_n)$ and suppose that $v \in V_x$ where $x \in V(K)$. If $x \in VW(K)$, then

$$d_p(v) = p \cdot (\lfloor \mu(x) \cdot n \rfloor - 1),$$

since v is connected by white edges to every $u \in V_x$ and by gray edges to every other vertex of G_n . Thus, for n sufficiently large, we have

$$\frac{d_p(v)}{|G_n|} = p \cdot \mu(x) + O(1/n) = g_K(p) + O(1/n),$$

thanks to part 3 of Proposition 2.4.

On the other hand, if $x \in VB(K)$, then v is connected by black edges to every $u \in V_x$ and by either gray or white edges to every other vertex of G_n , so

$$\begin{aligned} d_p(v) &= (1-p) \cdot (\lfloor \mu(x) \cdot n \rfloor - 1) + p \cdot \sum_{y \in V(K): xy \in EW(K)} \lfloor \mu(y) \cdot n \rfloor \\ &= (1-p) \cdot (\lfloor \mu(x) \cdot n \rfloor - 1) + p \cdot \left(|G_n| - \lfloor \mu(x) \cdot n \rfloor - \sum_{y \in V(K): xy \in EG(K)} \lfloor \mu(y) \cdot n \rfloor \right). \end{aligned}$$

Thus, by again appealing to part 3 of Proposition 2.4, for n sufficiently large, we have

$$\begin{aligned} \frac{d_p(v)}{|G_n|} &= (1-p)\mu(x) + p \cdot (1 - \mu(x) - d_G(x)) + O(1/n) \\ &= (1-p)\mu(x) + p \cdot \left(1 - \mu(x) - \frac{p - g_K(p)}{p} - \frac{1-2p}{p}\mu(x) \right) + O(1/n) \\ &= g_K(p) + O(1/n). \end{aligned}$$

We conclude that

$$\frac{\Delta_p(G_n)}{|G_n|} = g_K(p) + O(1/n),$$

and so the claim follows. $\square$

To go further, we consider modifying a CRG by replacing a white vertex by a collection of black vertices.

Definition 4.16. A *dalmatian CRG* of size ℓ is a CRG consisting of ℓ black vertices with all white edges.

Let K be a 0-core CRG with $|VW(K)| = w$, fix an integer $\ell \in \mathbb{N}$ and an integer $1 \leq r \leq w$. The CRG $K^\ell(r)$ is defined by replacing r of K 's white vertices by dalmatian CRGs of size ℓ and otherwise leaving K unchanged.

Dalmatian CRGs were introduced by Martin-Riasanovsky [10].

The following lemma expands on Proposition 4.12 and is the key step in the main result of this section. Intuitively, it states that if a collection of black vertices in a 0-core CRG behave a single white vertex (as per Proposition 4.12), then they actually behave like a dalmatian CRG.

Lemma 4.17. Let K, L be CRGs where K is 0-core and fix $n \in \mathbb{N}$. Let m be another positive integer and let $\mu \in \Delta^L$ be any probability mass satisfying $m \cdot \mu(x) \geq n$ for all $x \in VB(L)$.

If $|VW(K)| = s > t = |VW(L)|$ and $n \times K \subseteq L[\mu, m]$, then also $n \times K^n(r) \subseteq L[\mu, m]$ for any $1 \leq r \leq s - t$.

Proof. Observe that if $r > 1$, then $K^n(r) = (K^n(r-1))^n(1)$. Furthermore, $K^n(1)$ is still 0-core and $|VW(K^n(1))| = |VW(K)| - 1$. Thus, it suffices to establish the claim only for $r = 1$ since then the full claim follows by a straight-forward induction.

Let $W_1, \dots, W_s, B_1, \dots, B_k$ be the vertex partition of the colored graph $n \times K$ given by Definition 4.8, where the W_i 's correspond to the white vertices of K and the B_i 's correspond to the black vertices of K . Similarly, let $W'_1, \dots, W'_t, B'_1, \dots, B'_k$ be the vertex partition of the colored graph $L[\mu, m]$ given by Definition 4.14. Furthermore, let $\phi: V(n \times K) \rightarrow V(L[\mu, m])$ be an injection which exhibits $n \times K \subseteq L[\mu, m]$.

Fix any $i \neq j \in [s]$ and $\ell \in [t]$; since each edge between W_i and W_j is gray and all edges inside W'_ℓ are white, it cannot be the case that both $\phi(W_i) \cap W'_\ell$ and $\phi(W_j) \cap W'_\ell$ are nonempty. Since $s > t$, we may therefore suppose, without loss of generality, that $\phi(W_s) \subseteq \bigcup_{i=1}^k B'_i$. Similarly, all edges within W_s are white and all edges within each B'_i are black, so no two vertices of W_s can be mapped to the same B'_i . Thus, labeling $W_s = \{w_1, \dots, w_n\}$, and relabeling the B'_i 's if necessary, we may suppose that $\phi(w_i) \in B'_i$.

Now, again since K is 0-core, every edge incident to w_i is white or gray. Furthermore, since all edges within each B'_i are black, we observe that $\phi^{-1}(B'_i) = \{w_i\}$ for all $i \in [n]$. Additionally, since ϕ exhibits $n \times K \subseteq L[\mu, m]$, we know that all edges between B'_i and $\phi(v)$ for any $i \in [n]$ and any $v \in V(n \times K) \setminus W_s$ must be gray.

We now turn our attention to $n \times K^n(1)$, which we can consider to have vertex-partition

$$W_1, \dots, W_{s-1}, D_1, \dots, D_n, B_1, \dots, B_k,$$

where $D_1, \dots, D_n$ correspond to the black vertices in the new dalmatian set. Consider a function $\phi': V(n \times K^n(1)) \rightarrow V(L[\mu, m])$ with the following properties:

- $\phi'(v) = \phi(v)$ for all $v \in W_1 \cup \dots \cup W_{s-1} \cup B_1 \cup \dots \cup B_k$, and

- $\phi'(d) \in B'_i$ for all $i \in [n]$ and all $d \in D_i$.

By the assumption on m and μ , for all $i \in [n]$, we have $|B'_i| \geq n = |D_i|$, so such a ϕ' exists which is also an injection. Based on the properties of ϕ laid out above, this ϕ' realizes $n \times K^n(1) \subseteq L[\mu, m]$. $\square$

We are now ready to state and prove the main result in this section.

Theorem 4.18. *Let $\mathcal{H} = \text{Forb}(\mathcal{F})$ be a non-trivial hereditary property, fix $p \in (0, 1/2)$ and suppose that there is a sequence of CRGs $(\Gamma_n)_{n \in \mathbb{N}}$ with the following properties:*

- I. *There is a sequence $p_n \rightarrow p$ for which $\Gamma_n \in \mathcal{K}_{p_n}(\mathcal{H})$, and*
- II. *$\limsup_{n \rightarrow \infty} |VW(\Gamma_n)| = w$, and*
- III. *The Γ_n 's are pairwise distinct.*

Then there is a p -core CRG K with the following properties:

1. *$K \in \mathcal{K}_p(\mathcal{H})$, and*
2. *$g_K(p) \leq \limsup_{n \rightarrow \infty} g_{\Gamma_n}(p_n)$, and*
3. *$|VW(K)| = s > w$, and*
4. *For any integer $1 \leq r \leq s - w$ and any positive integer ℓ , $\mathcal{F} \not\vdash K^\ell(r)$.*

Proof. By passing to a subsequence of (Γ_n) , we can guarantee the following additional properties:

- IV. $p_n \in (0, 1/2)$, so $\Gamma_n \in \mathcal{K}_0(\mathcal{H})$, and (property I)
- V. $|VW(\Gamma_n)| = w$, and (property II and the fact that $w < \chi(\mathcal{F})$ is a non-negative integer)
- VI. $|VB(\Gamma_n)| \geq n$. (Properties III and V)

Let $\mu_n \in \Delta^{\Gamma_n}$ be the probability mass achieving $g_{\Gamma_n}(p_n)$. For each $n \in \mathbb{N}$, we can locate some integer N_n such that

- i. $\mu_n(x) \cdot N_n \geq n$ for all $x \in V(\Gamma_n)$, and (μ_n has full support)
- ii. $\left| g_{\Gamma_n}(p_n) - \frac{\Delta_{p_n}(\Gamma_n[\mu_n, N_n])}{|\Gamma_n[\mu_n, N_n]|} \right| < \frac{1}{n}$. (Proposition 4.15)

Set $G_n \stackrel{\text{def}}{=} \Gamma_n[\mu_n, N_n]$. Note that property i implies $n \times \Gamma_n \subseteq G_n$ and that property ii implies

$$\limsup_{n \rightarrow \infty} \frac{\Delta_{p_n}(G_n)}{|G_n|} = \limsup_{n \rightarrow \infty} g_{\Gamma_n}(p_n).$$

Furthermore,

$$|\Delta_{p_n}(G_n) - \Delta_p(G_n)| = \left| \max_{v \in V(G_n)} d_{p_n}(v) - \max_{v \in V(G_n)} d_p(v) \right| \leq \max_{v \in V(G_n)} |d_{p_n}(v) - d_p(v)| \leq |G_n| \cdot |p_n - p|,$$

so, since $p_n \rightarrow p$,

$$\limsup_{n \rightarrow \infty} \frac{\Delta_p(G_n)}{|G_n|} = \limsup_{n \rightarrow \infty} g_{\Gamma_n}(p_n).$$

We claim next that $K(w+1, 0) \sqsubset_{\rightarrow} (G_n)$. Indeed, Properties V and VI allow us to apply Proposition 4.12 to find that $n \times K(w+1, 0) \subseteq n \times \Gamma_n \subseteq G_n$, where the last inclusion follows from property i.

Thus, thanks to Lemma 4.13, we may find some p -core CRG K such that $|VW(K)| = s > w$ and $K \sqsubset_{\rightarrow} (G_n)$. Since $K \sqsubset_{\rightarrow} (G_n)$, we know that $\mathcal{F} \not\vdash K$; hence $K \in \mathcal{K}_p(\mathcal{H})$. Furthermore,

$$g_K(p) \leq \limsup_{n \rightarrow \infty} \frac{\Delta_p(G_n)}{|G_n|} = \limsup_{n \rightarrow \infty} g_{\Gamma_n}(p_n).$$

To finish the claim, fix any $1 \leq r \leq s - w$ and any $\ell \in \mathbb{N}$. In order to show that $\mathcal{F} \not\vdash K^\ell(r)$, it suffices to show that $K^\ell(r) \sqsubset_{\rightarrow} (G_n)$.

Fix any integer $N \geq \ell$. Since $K \sqsubset_{\rightarrow} (G_n)$, there must be some $n \geq N$ for which $N \times K \subseteq G_n$. Thus, property i allows us to apply Lemma 4.17 to see that also $N \times K^N(r) \subseteq G_n$. We conclude that

$$N \times K^\ell(r) \subseteq N \times K^N(r) \subseteq G_n,$$

and so $K^\ell(r) \sqsubset_{\rightarrow} (G_n)$ as needed. $\square$

We can now show that $\text{Forb}(K_{t,t})$ is free of accumulation points in $(0, 1]$.

Proof of Theorem 1.7. Set $\mathcal{H} = \text{Forb}(K_{t,t})$.

To begin, we note that $p = 1/2$ is never an accumulation point of any hereditary property due to [10, Theorem 39].

Next, observe that if $K \in \mathcal{K}_1(\mathcal{H})$, then K has at most one white vertex and fewer than $\overline{\chi}(K_{t,t}) = t$ black vertices. In particular, $\mathcal{K}_1(\mathcal{H})$ is finite and so $\mathcal{H}$ has no accumulation points in $(1/2, 1]$.

Now, fix $p \in (0, 1/2)$ and suppose for the sake of contradiction that p is an accumulation point of $\mathcal{H}$. Thus, we may find some sequence $p_n \rightarrow p$ and a sequence of CRGs $(\Gamma_n)_{n \in \mathbb{N}}$ for which

- $\Gamma_n \in \mathcal{K}_{p_n}(\mathcal{H})$, and
- The Γ_n 's are distinct, and
- $\text{ed}_{\mathcal{H}}(p_n) = g_{\Gamma_n}(p_n)$.

Since $\text{ed}_{\mathcal{H}}$ is continuous, this implies that

$$\lim_{n \rightarrow \infty} g_{\Gamma_n}(p_n) = \text{ed}_{\mathcal{H}}(p).$$

Thus, we may apply Theorem 4.18 to the sequence (Γ_n) . Let K be the p -core CRG guaranteed by Theorem 4.18 and note that $|VW(K)| \geq 1$. Since $K_{t,t}$ is bipartite and K has at least one white vertex, we know that K has exactly one white vertex and that $|VB(K)| \leq t - 1$. Furthermore, $K \in \mathcal{K}_p(\mathcal{H})$ and

$$g_K(p) \leq \limsup_{n \rightarrow \infty} g_{\Gamma_n}(p_n) = \text{ed}_{\mathcal{H}}(p) \implies g_K(p) = \text{ed}_{\mathcal{H}}(p).$$

Thus, Proposition 2.10 uniquely determines $K = K(1, t - 1)$. However, $K(1, t - 1)^1(1) = K(0, t)$ and, since $\overline{\chi}(K_{t,t}) = t$, we know that $K_{t,t} \mapsto K(0, t)$. This contradicts part 4 of Theorem 4.18 and so we have established the claim. $\square$

Unfortunately, we have been unable to rule out the possibility of other accumulation points for $\text{Forb}(K_{2,t})$, $t \geq 9$, though we believe that none exist.

Remark 4.19. Our proof of Theorem 1.7 actually establishes a stronger claim. Set $\mathcal{H} = \text{Forb}(K_{t,t})$, fix $p \in (0, 1]$ and fix any sequence $p_n \rightarrow p$. If there are CRGs $\Gamma_n \in \mathcal{K}_{p_n}(\mathcal{H})$ with $\lim_{n \rightarrow \infty} g_{\Gamma_n}(p_n) = \text{ed}_{\mathcal{H}}(p)$, then the sequence (Γ_n) is eventually constant.

5. Concluding remarks

While we gained a better understanding of accumulation points of the edit distance function in this paper, we still have many questions. For instance, for which p can a hereditary property have an accumulation point at p ?

Question 5.1. Let $\mathcal{A}$ be the set of all $p \in [0, 1/2]$ for which there is some non-trivial hereditary property with an accumulation point at p . What can be said about the set $\mathcal{A}$? In particular:

- What is $\sup \mathcal{A}$?
- Is $\mathcal{A}$ an interval?

Currently, all that we know about the set $\mathcal{A}$ is that $0, 1/4 \in \mathcal{A}$ and that $\sup \mathcal{A} \leq 1 - \varphi^{-1} = 0.382\dots$ where $\varphi = 1.618\dots$ is the golden ratio. The latter follows from the work of Martin and Riasanovsky [10, Theorem 39]. The work of Martin–Riasanovsky additionally suggests that perhaps $\sup \mathcal{A} \leq 1/3$ (cf. [10, Section 2.3 & Proposition 35]).

We expect that, at the very least, $1/n \in \mathcal{A}$ for all integers $n \geq 4$. Indeed, one could likely build on the ideas in the proof of Theorem 1.5 and construct a family $\mathcal{F}$ of $(n - 2)$ -regular graphs so that the property $\text{Forb}(\{K_{1,n}\} \cup \mathcal{F})$ has an accumulation point at $p = 1/n$.

Beyond this, it is worth pointing out that our construction of a hereditary property with an accumulation point at $p = 1/4$ from Theorem 1.5 required an infinite set of forbidden graphs. It is natural to wonder if this was necessary.

Question 5.2. Suppose that $\mathcal{H} = \text{Forb}(\mathcal{F})$ where $\mathcal{F}$ is finite. Can $\mathcal{H}$ have any accumulation points in $(0, 1)$?

We expect a negative answer, though, since this is the first systematic investigation into the accumulation points of the edit distance function, our current knowledge is severely limited. Regardless of the answer to the above question, we are confident to conjecture the following extension of Theorem 1.7.

Conjecture 5.3. For any $s, t \in \mathbb{N}$, $\text{Forb}(K_{s,t})$ has no accumulation points in the interval $(0, 1]$.

We note that it is necessary to go beyond the ideas used in Theorem 1.7 in order to establish this conjecture. Indeed, for every $k \geq 4$, Marchant–Thomason [6, Theorem 3.27] constructed infinitely many $1/k$ -core CRGs achieving $\text{ed}_{\text{Forb}(K_{1,k})}(1/k)$; therefore, the analogue of Remark 4.19 fails to hold in these cases.

Despite this, the same reasoning used in Theorem 1.7 implies that $\text{Forb}(K_{s,t})$ is free of accumulation points in $[1/2, 1]$ and that if $p \in (0, 1/2)$ happens to be an accumulation point, then $K(1, \min\{s, t\} - 1)$ achieves $\text{ed}_{\text{Forb}(K_{s,t})}(p)$.

We furthermore believe that accumulation points are relatively rare.

Conjecture 5.4. *Any non-trivial hereditary property has only finitely many accumulation points.*

Finally, we believe that accumulation points can occur only in one half of the interval $(0, 1)$.

Conjecture 5.5. *For any non-trivial hereditary property, either $(0, 1/2]$ or $[1/2, 1)$ is free of accumulation points.*

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] N. Alon, U. Stav, What is the furthest graph from a hereditary property?, *Random Struct. Algorithms* 33 (1) (2008) 87–104.
- [2] J. Balogh, R. Martin, Edit distance and its computation, *Electron. J. Comb.* 15 (1) (2008) 20.
- [3] B. Bollobás, A. Thomason, Hereditary and monotone properties of graphs, in: *The Mathematics of Paul Erdős, II*, in: *Algorithms Combin.*, vol. 14, Springer, Berlin, 1997, pp. 70–78.
- [4] B. Bollobás, A. Thomason, The structure of hereditary properties and colourings of random graphs, *Combinatorica* 20 (2) (2000) 173–202.
- [5] P. Erdős, A.H. Stone, On the structure of linear graphs, *Bull. Am. Math. Soc.* 52 (12) (1946) 1087–1091.
- [6] E. Marchant, A. Thomason, Extremal graphs and multigraphs with two weighted colours, in: *Fete of Combinatorics and Computer Science*, in: *Bolyai Soc. Math. Stud.*, vol. 20, János Bolyai Math. Soc., Budapest, 2010, pp. 239–286.
- [7] R. Martin, The edit distance function and symmetrization, *Electron. J. Comb.* 20 (3) (2013) 26.
- [8] R.R. Martin, The edit distance in graphs: methods, results, and generalizations, in: *Recent Trends in Combinatorics*, in: *IMA Vol. Math. Appl.*, vol. 159, Springer, Cham, 2016, pp. 31–62.
- [9] R.R. Martin, T. McKay, On the edit distance from $K_{2,t}$ -free graphs, *J. Graph Theory* 77 (2) (2014) 117–143.
- [10] R.R. Martin, A.W.N. Riasanovsky, On the edit distance function of the random graph, *Comb. Probab. Comput.* (2022), <https://doi.org/10.1017/S0963548321000353>, in press.
- [11] E. Szemerédi, Regular partitions of graphs, in: *Problèmes combinatoires et théorie des graphes*, *Colloq. Internat. CNRS*, Univ. Orsay, Orsay, 1976, in: *Colloq. Internat. CNRS*, vol. 260, CNRS, Paris, 1978, pp. 399–401.