Finding large expanders in graphs: from topological minors to induced subgraphs

Baptiste Louf
Centre de Recherches en Mathématique
Université de Montréal
Canada
baptiste.louf@umontreal.ca

Fiona Skerman*
Department of Mathematics
Uppsala University
Sweden
fiona.skerman@math.uu.se

Submitted: Nov 8, 2021; Accepted: Sep 27, 2022; Published: Jan 27, 2023

Abstract

In this paper, we consider a structural and geometric property of graphs, namely the presence of large expanders. The problem of finding such structures was first considered by Krivelevich [SIAM J. Disc. Math. 32 1 (2018)]. Here, we show that the problem of finding a large induced subgraph that is an expander can be reduced to the simpler problem of finding a large topological minor that is an expander. Our proof is constructive, which is helpful in an algorithmic setting.

We also show that every large subgraph of an expander graph contains a large subgraph which is itself an expander.

Mathematics Subject Classifications: 05C48, 05C83

1 Introduction

Expander graphs are important objects in graph theory. They have a wide variety of properties: for instance, the random walk mixes very fast [1, 6], and the eigenvalues of their Laplacian are well-separated [2, 4]. In computer science, they are used for clustering with the expander decomposition technique (see for instance [8]). We refer to [7] for an extensive survey of expander graphs and their applications.

Intuitively a graph is an edge-expander if there is a lower bound on the number of edges coming out of any vertex set relative to the 'size' of that vertex set. In this paper we define edge-expansion relative to the volume of the vertex set rather than its number of vertices: precise definitions for edge-expansion and conductance are given in Section 2.

*Partially supported by AI4Research at Uppsala University and by the Wallenberg AI, Autonomous Systems and Software Program WASP.

Unfortunately, the conditions for a graph to be an expander are quite restrictive, as the edge expansion of any subset of vertices should be high enough. In particular, the conductance of a graph is influenced by every "microscopic" part of the graph: for example adding a disjoint edge causes the conductance to drop to zero. There is a much larger set of graphs that shares macroscopic features with expanders and is more robust to small perturbations in the edge set. A natural such set is all graphs which contain linear sized expanders as induced subgraphs.

The problem of finding linear sized expander subgraphs has earlier been considered in [9, 10], where a sufficient local sparsity condition was given, with application to finding large minors, paths and cycles. This follows previous works on "weak" expander subgraphs in which the expansion factor is allowed to depend on the size of the graph [12, 13]. This question has also been studied in works posterior to ours, see e.g. [3].

However, finding an expander graph H which is an induced subgraph of G can be a difficult task, as the only operation that is permitted to find H in G is to iteratively delete a vertex and all its adjacent edges. Our main result is that this problem is actually equivalent to finding a topological minor that is an expander, which allows for more freedom in the operations that can be performed. More precisely, the following theorem holds.

Theorem 1. *For all $\kappa, \alpha > 0$, and for all $0 < \alpha' < \alpha$, there exists a $\kappa' > 0$ such that the following holds for every (multi)graph G.*

If there exists a graph H satisfying the following conditions:

- $e(H) \geqslant \alpha e(G)$,
- *H is a topological minor of G,*
- *H is a κ-expander,*

then there exists a graph H^* satisfying the following conditions:

- $e(H^*) \geqslant \alpha' e(G)$,
- *H^* is an induced subgraph of G,*
- *H^* is a κ'-expander.*

The proof is constructive, which implies that if we are given an algorithm to construct a topological minor of a graph that is an expander, we have an algorithm to construct an induced subgraph of that graph that is an expander. This result can be helpful to study some models of graphs where the topology is an important feature, such as *combinatorial maps*. In particular, Theorem 1 is a key tool in [11], which studies maps in the hyperbolic high genus regime, and the presence of large expander subgraphs in them. Our result is important for this problem as it then allows to use and study the classical "core-kernel decomposition", which produces topological minors of maps.

The proof of the main theorem strongly relies on the following theorem of independent interest, that states that every large subgraph of an expander contains a large subgraph that is itself an expander. We mention that this next result is related to Theorem 2.7 of [10]. There, a very mild condition on G - expansion of certain sized subsets of G - is shown to imply a linear sized induced subgraph G^* which is an expander.

Theorem 2. *Take $0 < \kappa \leqslant 1$ and $0 < \varepsilon < \frac{\kappa}{6}$, and G a κ-expander. Then for any subgraph H of G satisfying $e(H) \geqslant (1-\varepsilon)e(G)$, there exists an induced subgraph H^* of H satisfying*

- $e(H^*) \geqslant \left(1 - \frac{6\varepsilon}{\kappa}\right) e(G)$,
- *H^* is a $\frac{\kappa}{3}$-expander.*

Structure of the paper We start with some definitions (Section 2), then in Section 4, we explain how Theorem 1 boils down to two propositions whose proofs are to be found in Sections 5 and 6. Section 3 contains the proof of Theorem 2. We prove Theorem 2 first as it will be used in the proof of the second proposition which comprises Theorem 1.

2 Definitions

A *graph* G is the data of a set of *vertices* $V(G)$ and a multiset $E(G)$ of *edges* whose elements are unordered pairs of (non-necessarily distinct) vertices in $V(G)$. Note that our definition of a graph allows loops and multiple edges, it is also common to call these objects multigraphs. For any graph G, we write $e(G)$ for its number of edges. We also write $e_G(X)$ and $e_G(X,Y)$ for the number of edges in X and the number of edges between vertex sets X and Y respectively. If $X \subset V(G)$, then we write $\overline{X} := V(G) \setminus X$. The *degree* of a vertex is the number of non-loop edges it belongs to, plus twice the number of loops it belongs to. If $X \subset V(G)$, then we write $\mathrm{vol}(X) = \sum_{v \in X} \deg(v)$. Given a vertex v of degree 2 not incident to a loop, we call *smoothing* v the operation that consists in replacing v and its two incident edges by a single edge (see Figure 1).

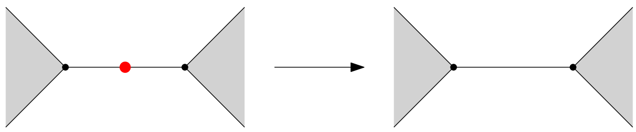

Figure 1: Smoothing a vertex of degree 2.

If G is a graph, an *induced subgraph* of G is a graph obtained from G by deleting some of its vertices. A *subgraph* of G is a graph obtained from G by deleting some of its vertices and edges. A *topological minor* of G is a graph obtained from G by deleting some of its vertices and edges, and smoothing some of its vertices. For H a subgraph of G we write $G \setminus H$ to denote the graph with vertex set $V(G)$ and edge set $E(G) \setminus E(H)$.

There are multiple notions of edge-expansion. It will be helpful for us to define edge-expansion as a function of edges out of a set normalised by the volume of the set, as in [5], rather than normalised by the number of vertices in the set, as [10]. Note these notions are essentially equivalent for regular graphs or graphs of bounded degree.

The *edge-expansion* of the vertex set $X \subset V(G)$ is defined as

$$h_G(X) := \frac{e_G(X, \overline{X})}{\min(\mathrm{vol}_G(X), \mathrm{vol}_G(\overline{X}))}$$

where we recall that $e_G(X, \overline{X})$ is the number of edges of G that have one endpoint in X and one endpoint in $\overline{X}$. The graph G is said to be a *κ-expander* if, for all $X \subset V(G)$, we have

$$h_G(X) \geqslant \kappa.$$

Notice that we actually only have to check the condition above for subsets X such that $G[X]$ is connected where $G[X]$ is the induced subgraph of G obtained by deleting all vertices in $\overline{X}$.

The *conductance* of a graph G is the largest κ such that G is a κ-expander. This is also sometimes called the Cheeger constant of the graph. However the Cheeger constant can also refer to the corresponding notion where $h'(G)$ is defined as $h(G)$ but normalised not by the volume but by the number of vertices; so we use the term conductance to avoid confusion.

Finally we make some definitions that are specific to this paper. We say a *vertex-coloured graph* is a graph that has a certain number of its vertices of degree 2 coloured in blue. A blue path of size ℓ is a path of ℓ consecutive blue vertices. If G is a vertex-coloured graph, then we call $\mathrm{red}(G)$ (the "reduced" version of G) the graph obtained from G by smoothing all its blue vertices.

An *edge-coloured graph* is a graph that has a certain number of edges coloured in blue. If G is an edge-coloured graph, then we call $\mathrm{red}(G)$ the graph obtained from G by deleting all its blue edges. If G is an edge-coloured graph, and $v \in V(G)$, then we write $d_{\mathrm{red}}(v)$ for the degree of v in $\mathrm{red}(G)$.

3 Proof of Theorem 2

This section proves Theorem 2 which will be used in the proof of Proposition 5 which is in turn used to prove Theorem 1. We start with an easy technical lemma that will be useful in the future.

Lemma 3. *Let H a graph and H' an induced subgraph of H. Let $X \subset V(H')$, we can also consider X as a subset of $V(H)$. If*

$$vol_{H'}(X) \leqslant vol_{H'}(\overline{X})$$

then

$$vol_H(X) \leqslant vol_H(\overline{X}).$$

Proof. Note that in any graph G,

$$\mathrm{vol}_G(X) \leqslant \mathrm{vol}_G(\overline{X})$$

is equivalent to

$$e_G(X) \leqslant e_G(\overline{X}).$$

Hence we may suppose $e_{H'}(X) \leqslant e_{H'}(\overline{X})$ and it remains only to show $e_H(X) \leqslant e_H(\overline{X})$. Since $X \subset V(H')$ and H' is an induced subgraph of H, we have

$$e_{H'}(X) = e_H(X).$$

On the other hand

$$e_{H'}(\overline{X}) = e_{H'}(V(H') \setminus X) = e_H(V(H') \setminus X)$$

where the last equality holds because H' is an induced subgraph of H. We also have

$$e_H(\overline{X}) = e_H(V(H) \setminus X).$$

Since $V(H') \setminus X \subset V(H) \setminus X$, we have

$$e_H(X) = e_{H'}(X) \leqslant e_{H'}(\overline{X}) = e_H(V(H') \setminus X) \leqslant e_H(V(H) \setminus X) = e_H(\overline{X}),$$

which concludes the proof. □

We recall the assumptions of Theorem 2. We are given G that is a κ-expander, and a subgraph H of G satisfying $e(H) \geqslant (1-\varepsilon)e(G)$. We are looking for a big enough induced subgraph H^* of H that is a $\frac{\kappa}{3}$-expander.

The strategy is the following: if H itself is not a $\frac{\kappa}{3}$-expander, we remove from it a "bad set" of vertices that has a low edge expansion. If we do not obtain a $\frac{\kappa}{3}$-expander, we go on and remove another bad set, and we keep going until we have a $\frac{\kappa}{3}$-expander. We will show that the process terminates and that the total size of the bad sets is small.

Setup We will construct a sequence $X_1, X_2, \dots$ of disjoint subsets of $V(H)$. We first introduce some notation: $Y_j = \bigcup_{i \leqslant j} X_i$, $H_j = H[V(H) \setminus Y_{j-1}]$. We will write $\mathrm{down}(X_i) = e_H(X_i, V(H_{i+1}))$, $\mathrm{diff}(X_i) = e_G(X_i) - e_H(X_i)$ and $\mathrm{out}(X_i) = e_{G \setminus H}(X_i, V(G) \setminus X_i)$, i.e. $\mathrm{out}(X_i)$ will be the number of edges of G that do not belong to H with exactly one endpoint in X_i. We also write

$$\Delta_j = \sum_{i \leqslant j-1} \mathrm{diff}(X_i)$$

and

$$O_j = \sum_{i \leqslant j-1} \mathrm{out}(X_i).$$

See Figure 2 for an illustration of the process.

The process Now we can define our process. At step i, if H_i is a $\frac{\kappa}{3}$-expander, then stop the process and output H_i. Otherwise, there must exist a nonempty subset X_i of $V(H_i)$ with

$$\mathrm{vol}_{H_i}(X_i) \leqslant \mathrm{vol}_{H_i}(\overline{X}_i)$$

such that

$$\mathrm{down}(X_i) < \frac{\kappa}{3}\mathrm{vol}_{H_i}(X_i) \leqslant \frac{\kappa}{3}\mathrm{vol}_G(X_i). \tag{1}$$

It is immediate this process terminates because H_i gets strictly smaller at each step.

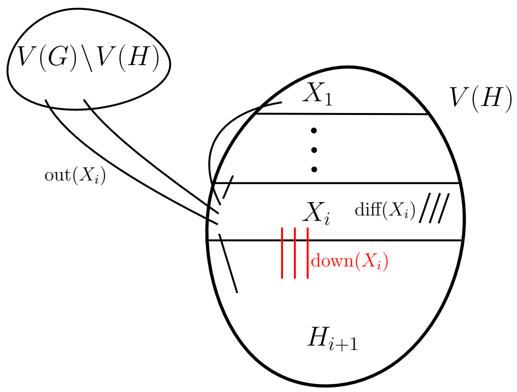

Figure 2: Illustration of the process: H is a subgraph of G with red edges depicting those in H and black edges those in $G \setminus H$.

Lower bounding the size of the output Fix some j. First, note that O_j counts edges that are in G but not in H (each edge can be counted twice), and same for Δ_j (each edge is only counted once this time). The edges counted by O_j and Δ_j are disjoint, hence

$$O_j/2 + \Delta_j \leqslant \varepsilon e(G). \tag{2}$$

Also, notice that

$$e_G(X_i, \overline{X}_i) = \mathrm{down}(X_i) + \sum_{k<i} e_H(X_i, X_k) + \mathrm{out}(X_i). \tag{3}$$

For all i, H_i is an induced subgraph of H, hence by Lemma 3 we have $\mathrm{vol}_H(X_i) \leqslant \mathrm{vol}_H(\overline{X}_i)$, and therefore $\mathrm{vol}_G(X_i) - 2\mathrm{diff}(X_i) \leqslant \mathrm{vol}_G(\overline{X}_i)$. Since G is a κ-expander we have

$$\kappa(\mathrm{vol}_G(X_i) - 2\mathrm{diff}(X_i)) \leqslant e_G(X_i, \overline{X}_i)$$

which implies, since $\kappa \leqslant 1$,

$$\kappa \mathrm{vol}_G(X_i) \leqslant e_G(X_i, \overline{X}_i) + 2\mathrm{diff}(X_i). \tag{4}$$

Observe by (3) and (4) respectively

$$\begin{aligned} \sum_{k<i} e_H(X_k, X_i) + \mathrm{out}(X_i) + 2\mathrm{diff}(X_i) &= e_G(X_i, \overline{X}_i) - \mathrm{down}(X_i) + 2\mathrm{diff}(X_i) \\ &\geqslant \kappa \mathrm{vol}_G(X_i) - \mathrm{down}(X_i). \end{aligned}$$

Hence by (1) applied twice

$$\sum_{k<i} e_H(X_k, X_i) + \mathrm{out}(X_i) + 2\mathrm{diff}(X_i) \geqslant \frac{2\kappa}{3}\mathrm{vol}_G(X_i) \geqslant 2\mathrm{down}(X_i).$$

Summing over all $i \leqslant j-1$, we obtain

$$\sum_{k<i\leqslant j-1} e_H(X_k, X_i) + O_j + 2\Delta_j \geqslant \frac{2\kappa}{3}\mathrm{vol}_G(Y_{j-1}) \geqslant 2\sum_{i\leqslant j-1}\mathrm{down}(X_i). \tag{5}$$

Now, for all fixed k, note that

$$\sum_{k<i} e_H(X_k, X_i) \leqslant \mathrm{down}(X_k),$$

hence the first inequality of (5) implies

$$\frac{2\kappa}{3}\mathrm{vol}_G(Y_{j-1}) \leqslant \sum_{k\leqslant j-2}\mathrm{down}(X_k) + O_j + 2\Delta_j, \tag{6}$$

and the second inequality of (5) implies

$$O_j + 2\Delta_j \geqslant 2\sum_{k\leqslant j-1}\mathrm{down}(X_k) - \sum_{k\leqslant j-2}\mathrm{down}(X_k) \geqslant \sum_{k\leqslant j-2}\mathrm{down}(X_k). \tag{7}$$

Combining (6) and (7), we obtain

$$\mathrm{vol}_G(Y_{j-1}) \leqslant \frac{3}{\kappa}(O_j + 2\Delta_j) \leqslant \frac{6}{\kappa}\varepsilon e(G) \tag{8}$$

where the second inequality comes from (2).

Now we can prove Theorem 2:

Proof of Theorem 2. Take G and H satisfying the assumptions of the theorem. Now, run the process until it stops at time j and outputs H_j. Take $H^* = H_j$. It is a $\frac{\kappa}{3}$-expander by definition of the process. We have $e(H^*) \geqslant e(G) - \mathrm{vol}_G(Y_{j-1})$, which by (8) finishes the proof. □

4 Strategy of the proof of Theorem 1

To obtain an expander induced subgraph from an expander topological minor, we proceed in two steps: first we find an expander subgraph of roughly the same size as the topological minor, then we find an expander induced subgraph of roughly the same size as this subgraph. Theorem 1 is an immediate corollary of the following two propositions.

Proposition 4. *For all $\kappa, \alpha > 0$, and for all $0 < \alpha' < \alpha$, there exists a $\kappa' > 0$ such that the following holds for every graph G.*

If there exists a graph H satisfying the following conditions:

- $e(H) \geqslant \alpha e(G)$,

- *H is a topological minor of G,*
- *H is a κ-expander,*

then there exists a graph H^* satisfying the following conditions:

- *$e(H^*) \geqslant \alpha' e(G)$,*
- *H^* is a subgraph of G,*
- *H^* is a κ'-expander.*

Proposition 5. *For all $\kappa, \alpha > 0$, and for all $0 < \alpha' < \alpha$, there exists a $\kappa' > 0$ such that the following holds for every graph G.*

If there exists a graph H satisfying the following conditions:

- *$e(H) \geqslant \alpha e(G)$,*
- *H is a subgraph of G,*
- *H is a κ-expander,*

then there exists a graph H^* satisfying the following conditions:

- *$e(H^*) \geqslant \alpha' e(G)$,*
- *H^* is an induced subgraph of G,*
- *H^* is a κ'-expander.*

5 From topological minors to subgraphs

In this section, we prove Proposition 4. We start by proving that replacing edges by paths of bounded length in an expander still yields an expander.

Lemma 6. *Let $M \geqslant 2$ and let H be a κ-expander. Suppose G is obtained from H by replacing each edge by a path with at most M edges. Then G is a $\frac{\kappa}{2M-1}$-expander.*

Proof. In G, colour in red the vertices that come from H, and the rest in black. Let Y be a subset of $V(G)$ such that $G[Y]$ is connected. Let X be the set of red vertices in Y. We want to lower bound $h_G(Y)$ in terms of $h_H(X)$. See Figure 3 for an illustration.

If $X = \emptyset$, then $G[Y]$ is a path on $\leqslant M - 1$ vertices and so $\mathrm{vol}_G(Y) \leqslant 2(M-1)$ and $e_G(Y, \overline{Y}) = 2$. Hence

$$h_G(Y) \geqslant \frac{e_G(Y, \overline{Y})}{\mathrm{vol}_G(Y)} \geqslant \frac{1}{M-1}.$$

Similarly if $\overline{X} = \emptyset$ then $h_G(Y) = h_G(\overline{Y}) \geqslant 1/(M-1)$.

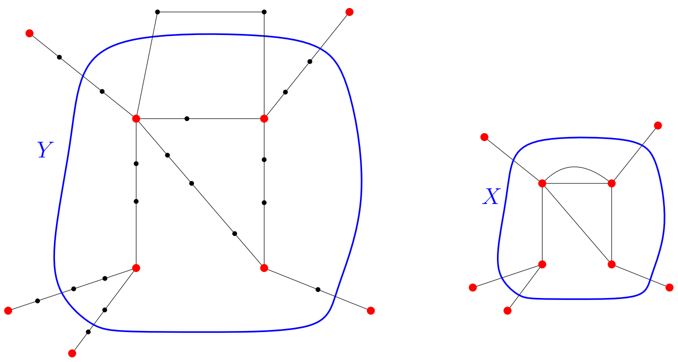

Figure 3: Comparing the edge expansions of Y in G and X in H in Lemma 6. Here, $M = 3$.

Let us consider now the case $X \neq \emptyset$ and $\overline{X} \neq \emptyset$. The number of edges of H which are incident to a vertex of X is $e_H(X) + e_H(X, \overline{X}) \leqslant \mathrm{vol}_H(X)$. Each edge of H is replaced by a path with at most M edges, therefore the number of black vertices in Y can be bounded above by $(M-1)\mathrm{vol}_H(X)$. All these vertices have degree 2, by definition of G, hence

$$\mathrm{vol}_G(Y) \leqslant \mathrm{vol}_H(X) + 2(M-1)\mathrm{vol}_H(X) = (2M-1)\mathrm{vol}_H(X) \tag{9}$$

and similarly

$$\mathrm{vol}_G(\overline{Y}) \leqslant (2M-1)\mathrm{vol}_H(\overline{X}). \tag{10}$$

Now, each edge counted in $e_H(X, \overline{X})$ corresponds to a path in G between Y and $\overline{Y}$, therefore

$$e_G(Y, \overline{Y}) \geqslant e_H(X, \overline{X}). \tag{11}$$

It is easy to check that $a' \leqslant ca$ and $b' \leqslant cb$ together imply $\min\{a', b'\} \leqslant c \min\{a, b\}$ and therefore, by (9), (10) and (11):

$$h_G(Y) \geqslant \frac{1}{2M-1} h_H(X).$$

This concludes the proof. □

The next proposition helps us find large subgraphs without long blue paths in vertex-coloured graphs. Recall a vertex coloured graph G has a subset of its degree two vertices coloured blue and red(G) is obtained from G by smoothing all its blue vertices.

Proposition 7. *For all $\kappa, \varepsilon, \alpha > 0$, there exists M such that the following holds for any vertex-coloured graph G such that $red(G)$ is a κ-expander and $e(red(G)) \geqslant \alpha e(G)$. If G_M is the induced subgraph of G obtained by deleting all the blue vertices in blue paths of length greater than M, and C_M is the connected component of G_M such that $e(red(C_M))$ is maximal, then*

$$e(red(C_M)) \geqslant (1-\varepsilon)e(red(G)),$$

and $red(C_M)$ is a subgraph of $red(G)$.

Note that $\mathrm{red}(C_M)$ need not be an induced subgraph of $\mathrm{red}(G)$. A path in G of blue vertices of length greater than M from C_M to itself will form an edge in $\mathrm{red}(G)$ but not in $\mathrm{red}(C_M)$.

Proof. The last point of the proposition follows from the fact that C_M is a subgraph of G. See Figure 4 for an illustration of this proof.

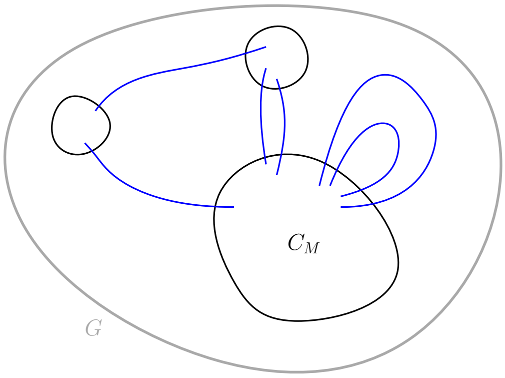

Figure 4: In black, the components of G_M, in blue, the blue paths of length $> M$, as considered in Proposition 7.

Let M be the smallest integer satisfying

$$M > \frac{\alpha}{1-\alpha}\frac{1+\frac{1}{\kappa}}{\varepsilon}. \tag{12}$$

Let p_M be the number of blue paths of size $> M$ in G. The quantity $e(G) - e(\mathrm{red}(G))$ counts exactly the number of blue vertices in G, hence

$$Mp_M \leqslant e(G) - e(\mathrm{red}(G)).$$

Now, using the inequality $e(G) \leqslant \frac{1}{\alpha}e(\mathrm{red}(G))$, we obtain

$$Mp_M \leqslant \frac{1-\alpha}{\alpha}e(\mathrm{red}(G))$$

which by (12) implies

$$p_M \leqslant \frac{\varepsilon}{1+\frac{1}{\kappa}}e(\mathrm{red}(G)). \tag{13}$$

We also have

$$e(\mathrm{red}(G)) = e(\mathrm{red}(G_M)) + p_M. \tag{14}$$

Let $\mathcal{C}(G_M)$ denote the set of connected components in G_M. Notice that C being a subgraph of G implies that $\mathrm{red}(C)$ is a subgraph of $\mathrm{red}(G)$, thus $\overline{\mathrm{red}(C)} = \mathrm{red}(G) \setminus \mathrm{red}(C)$ is well-defined. Consider $C \in \mathcal{C}(G_M)$. In $\mathrm{red}(G)$, any edge connecting $\mathrm{red}(C)$ to $\overline{\mathrm{red}(C)}$ corresponds to a blue path of size $> M$ in G. Hence

$$\sum_{C\in\mathcal{C}(G_M)} e_{\mathrm{red}(G)}(\mathrm{red}(C), \overline{\mathrm{red}(C)}) \leqslant 2p_M \tag{15}$$

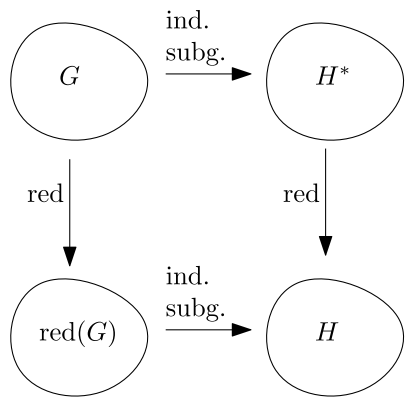

Figure 5: Building H^* out of H in the proof of Lemma 8. An "ind. subg." arrow means that the target is an induced subgraph of the source. A "red" means that the target is obtained from the source by applying the operation red.

(this is not an equality since some blue paths of size $> M$ might go from a component $C \in \mathcal{C}(G_M)$ to itself, see Figure 4).

For every connected component $C \neq C_M$ of G_M, we have

$$e(\mathrm{red}(C)) \leqslant e(\overline{\mathrm{red}(C)}),$$

and since $\mathrm{red}(G)$ is a κ-expander, we have:

$$e_{\mathrm{red}(G)}(\mathrm{red}(C), \overline{\mathrm{red}(C)}) \geqslant 2\kappa e(\mathrm{red}(C))$$

because the volume in $\mathrm{red}(G)$ of the vertices of $\mathrm{red}(C)$ is bigger than $2e(\mathrm{red}(C))$. If we sum over all $C \neq C_M$, we have, by (15)

$$p_M \geqslant \kappa(e(\mathrm{red}(G_M)) - e(\mathrm{red}(C_M)))$$

therefore, by (14)

$$e(\mathrm{red}(C_M)) \geqslant e(\mathrm{red}(G)) - (1 + \frac{1}{\kappa})p_M$$

and because of (13) we obtain

$$e(\mathrm{red}(C_M)) \geqslant (1 - \varepsilon)\, e(\mathrm{red}(G)),$$

which concludes the proof. □

The following lemma helps us invert the red operation in some sense.

Lemma 8. *Let G be a vertex-colored graph, and H an induced subgraph of red(G), then there exists a vertex-colored graph H^* that is an induced subgraph of G such that red(H^*) = H.*

Proof. See Figure 5 for visual aid. Let $X = V(G) \setminus V(H)$. This set can be seen as a subset of $V(G)$. Let us build a set $Y \subset V(G)$ from X in the following way:

- a black vertex belongs to Y iff it belongs to X,
- Notice that every blue vertex is part of a blue path that is incident to two black vertices. A blue vertex belongs to Y iff at least one of these two black vertices belongs to X.

Now, let H^* be the graph obtained from G by deleting all vertices of Y. We claim that H^* is a vertex-colored graph and that $\text{red}(H^*) = H$. To show that H^* is vertex colored, we need to show that all its blue vertices have degree 2, or equivalently that any blue vertex of G that is incident to a vertex of Y is also in Y (loosely speaking, if you delete a neighbor of a blue vertex, you also delete this vertex). This last condition is ensured by the construction of Y. Finally, $\text{red}(H^*) = H$ is immediate because $V(H) \subset V(H^*)$ and $V(H^*) \setminus V(H)$ only contains blue vertices (again, here we see $V(H)$ as a subset of $V(G)$). ☐

Finally, after recalling the following fact, we are ready to prove Proposition 4.

Lemma 9 (folklore)**.** *Any subgraph of a graph can be obtained by first applying vertex deletions, then edge deletions. Any topological minor of a graph can be obtained by first applying vertex deletions, then edge deletions, then smoothings.*

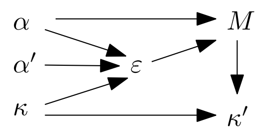

Figure 6: The causal graph of the variables we introduced in the proof of Proposition 4. Note that α, α' and κ are the constants introduced in the statement of the proposition with the claim that $\kappa' = \kappa'(\alpha, \alpha', \kappa)$.

Proof of Proposition 4. This proof might be complicated to follow, Figure 7 presents the relations between the different graphs involved. Also, we introduce a lot of variables, Figure 6 presents a causal graph of all the variables to make sure there is no circularity. Let G and H be graphs satisfying the assumptions of the proposition. Let $\varepsilon > 0$ be such that

$$\alpha\left(1 - \frac{6\varepsilon}{\kappa}\right) = \alpha'. \tag{16}$$

By Lemma 9, there must exist a vertex-coloured graph G_1 such that G_1 is a subgraph of G (if we forget about the colouring), and such that $\text{red}(G_1) = H$.

Now, by Proposition 7, there exists M depending only on κ, α and ε, and a subgraph C_M of G_1 such that $e(\text{red}(C_M)) \geqslant (1-\varepsilon)e(H)$ and such that C_M does not contain any blue path of length $> M$.

Hence we have $\text{red}(C_M)$ a subgraph of H, with $e(\text{red}(C_M)) \geqslant (1-\varepsilon)e(H)$ and H a κ-expander. Thus by Theorem 2 and (16) in the proof of Proposition 4, there exists an induced subgraph H' of $\text{red}(C_M)$, such that $e(H') \geqslant \alpha' e(G)$, such that H' is a $\frac{\kappa}{3}$-expander.

Now, by Lemma 8, there exists a graph H^* that is a subgraph of C_M such that $\text{red}(H^*) = H'$. It is direct to see that H^* does not contain any blue path of length $> M$. Therefore, by Lemma 6, H^* is a κ'-expander, where

$$\kappa' = \frac{\kappa}{3(2M-1)}.$$

Furthermore, we have $e(H^*) \geqslant e(H') \geqslant \alpha' e(G)$, which concludes the proof. □

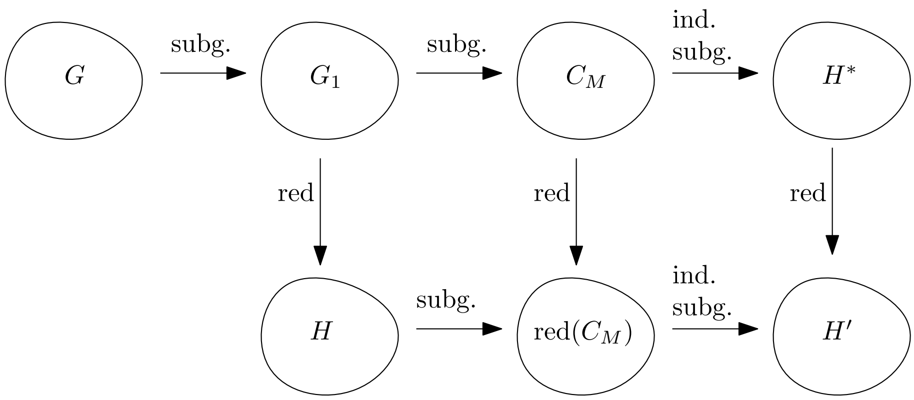

Figure 7: The relations between the different graphs considered in the proof of Proposition 4. An "ind. subg." arrow means that the target is an induced subgraph of the source. A "subg." arrow means that the target is a subgraph of the source. A "red" means that the target is obtained from the source by applying the operation red.

6 From subgraph to induced subgraph

Here we prove Proposition 5. The structure of the proof is similar to the proof of Proposition 4. However, it is a bit more technical as the arguments are more intertwined than in the previous section.

The following lemma is similar to Lemma 6, but for adding edges instead of paths.

Lemma 10. *Let H be a κ-expander, and let G be obtained from H adding some edges between its vertices, such that no vertex sees its degree multiplied by more than M. Then G is a $\frac{\kappa}{M}$-expander.*

Proof. We stress the fact that $V(G) = V(H)$. For all $X \subset V(G)$, we have immediately

$$\text{vol}_G(X) \leqslant M \text{vol}_H(X)$$

and

$$e_G(X, \overline{X}) \geqslant e_H(X, \overline{X}),$$

therefore the lemma follows.

□

Recall an edge-coloured graph G has a subset of its edges coloured blue, $\mathrm{red}(G)$ is obtained from G by deleting all its blue edges and for $v \in V(G)$ we denote by $d_{\mathrm{red}}(v)$ the degree of v in $\mathrm{red}(G)$.

Lemma 11. *For all $\varepsilon, \alpha > 0$, there exists M such that the following holds for any edge-coloured graph G such that $e(red(G)) \geqslant \alpha e(G)$. If $V^* = \{v \in V(G) \mid \deg(v) \geqslant M d_{red}(v)\}$, then*

$$vol_{red(G)}(V^*) \leqslant \varepsilon e(red(G)).$$

Proof. This is immediate because

$$\mathrm{vol}_{\mathrm{red}(G)}(V^*) \leqslant \frac{1}{M}\mathrm{vol}_G(V^*) \leqslant \frac{1}{M}e(G) \leqslant \frac{1}{M\alpha}e(\mathrm{red}(G)).$$ □

The following proposition will play the same role as Theorem 2 and Proposition 7 in Section 5. Unfortunately, this time we cannot make it into two separate arguments, which makes things a little more technical.

Proposition 12. *For all $\kappa, \varepsilon, \alpha > 0$, there exists M such that the following holds for any edge-coloured graph G. If $red(G)$ is a κ-expander and $e(red(G)) \geqslant \alpha e(G)$, then there exists an induced subgraph G^* of G such that $red(G^*)$ is a $\frac{\kappa}{3}$-expander, and for every vertex v of G^*, $\deg(v) \leqslant 3M d_{red}(v)$, and $e(G^*) \geqslant \left(1 - \varepsilon\left(1 + \frac{3}{\kappa}\right)\right) e(red(G))$.*

Proof. We will construct an algorithm which takes as input a graph G satisfying the assumptions of the proposition, and outputs a graph G^* with the required properties.

Setup Some notation before we start: we will construct a finite sequence $X_0, X_1, \dots$ of disjoint subsets of $V(G)$. We will write $Y_j = \bigcup_{0 \leqslant i \leqslant j} X_i$, $V_j = V(G) \setminus Y_{j-1}$ and $G_j = G[V_j]$. We will also write $H = \mathrm{red}(G)$ and $H_j = \mathrm{red}(G_j)$.

Finally, let $\mathrm{up}(X_i) = e_H(X_i, Y_{i-1})$ and $\mathrm{down}(X_i) = e_H(X_i, V_{i+1}) = e_{H_i}(X_i, V_{i+1})$ (the second equality is a consequence of the definitions). See Figure 8 for an illustration of the process.

The algorithm The algorithm works this way:

At step 0, let V^* and M be given by Lemma 11. Set $X_0 := V^*$.

At step $i > 0$:

Case 1: *If* there exists a set $X \subset V_i$ such that

$$\mathrm{vol}_{G_i}(X) \geqslant 3M\mathrm{vol}_{H_i}(X),$$

set $X_i := X$.

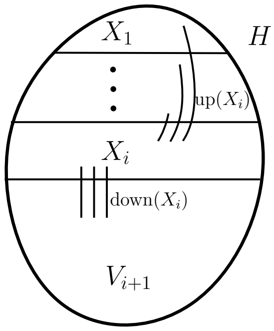

Figure 8: Illustration of the process in the proof of Proposition 12.

Case 2: *Else, if* there exists a set $X \subset V_i$ such that $\text{vol}_{H_i}(X) \leqslant \text{vol}_{H_i}(\overline{X})$ and

$$e_{H_i}(X, \overline{X}) \leqslant \frac{\kappa}{3}\text{vol}_{H_i}(X),$$

set $X_i := X$.

Else, return G_i.

Note that this algorithm terminates because G_i gets strictly smaller at each step.

Estimating the size of the output graph: We let the algorithm run until it stops and returns some G_j. Let us analyse each step of the algorithm. Take a step $i > 0$.

If we are in case 1, we have

$$\text{vol}_{G_i}(X_i) \geqslant 3M\text{vol}_{H_i}(X_i)$$

but

$$\text{vol}_G(X_i) \leqslant M\text{vol}_H(X_i)$$

by definition of X_0.

Hence, knowing that $\text{vol}_H(X_i) = \text{up}(X_i) + \text{vol}_{H_i}(X_i)$, and $\text{vol}_{G_i}(X_i) \leqslant \text{vol}_G(X_i)$ we obtain the following inequality

$$\text{up}(X_i) \geqslant \frac{2}{3}\text{vol}_H(X_i).$$

Using the fact $\text{vol}_{H_i}(X_i) \geqslant \text{down}(X_i)$, we can obtain the following inequality

$$\text{up}(X_i) \geqslant 2\text{down}(X_i).$$

If we are in case 2, we have

$$\text{down}(X_i) \leqslant \frac{\kappa}{3}\text{vol}_{H_i}(X_i)$$

but, H is a κ-expander, and by Lemma 3 $\text{vol}_{H_i}(X_i) \leqslant \text{vol}_{H_i}(\overline{X}_i)$ implies $\text{vol}_H(X_i) \leqslant \text{vol}_H(\overline{X}_i)$, hence

$$\text{up}(X_i) + \text{down}(X_i) \geqslant \kappa\text{vol}_H(X_i).$$

Since $\mathrm{vol}_{H_i}(X_i) \leqslant \mathrm{vol}_H(X_i)$, this implies immediately

$$\mathrm{up}(X_i) \geqslant \frac{2\kappa}{3}\mathrm{vol}_H(X_i)$$

and

$$\mathrm{up}(X_i) \geqslant 2\mathrm{down}(X_i).$$

In any case, we have for all $i > 0$

$$\mathrm{up}(X_i) \geqslant 2\mathrm{down}(X_i), \tag{17}$$

and since $\kappa \leqslant 1$,

$$\mathrm{up}(X_i) \geqslant \frac{2\kappa}{3}\mathrm{vol}_H(X_i). \tag{18}$$

Finally, let

$$S_j = \sum_{0<i\leqslant j-1} \mathrm{up}(X_i).$$

By definition of up and down, we have

$$S_j \leqslant \sum_{0\leqslant i\leqslant j-2} \mathrm{down}(X_i),$$

which, by (17) implies

$$S_j \leqslant \mathrm{down}(X_0) + \frac{1}{2}S_{j-1}.$$

Using the trivial inequalities $S_{j-1} \leqslant S_j$ and $\mathrm{down}(X_0) \leqslant \mathrm{vol}_H(X_0)$, we obtain

$$S_j \leqslant 2\mathrm{vol}_H(X_0). \tag{19}$$

On the other hand, by summing (18) over $0 < i \leqslant j-1$, we have for $j \geqslant 2$

$$S_j \geqslant \frac{2\kappa}{3}\Big(\mathrm{vol}_H(Y_{j-1}) - \mathrm{vol}_H(X_0)\Big). \tag{20}$$

Combining (19) and (20), one obtains for $j \geqslant 2$

$$\mathrm{vol}_H(Y_{j-1}) \leqslant \left(1 + \frac{3}{\kappa}\right)\mathrm{vol}_H(X_0). \tag{21}$$

Notice that if $j = 1$ then $Y_{j-1} = X_0$ and so (21) holds for all j. By Lemma 11, we have $\mathrm{vol}_H(X_0) \leqslant \varepsilon e(H)$. Now, since $e(G_j) \geqslant e(H_j) \geqslant e(H) - \mathrm{vol}_H(Y_{j-1})$, by (21) we have

$$e(G_j) \geqslant \left(1 - \varepsilon\left(1 + \frac{3}{\kappa}\right)\right)e(H).$$

The condition of case 1 ensures that for all v of G_j, $\deg(v) \leqslant 3Md_{\mathrm{red}}(v)$, and the condition of case 2 ensures that $\mathrm{red}(G_j)$ is a $\frac{\kappa}{3}$-expander. We set $G^* = G_j$, this finishes the proof. □

We can now turn to the proof of Proposition 5

Proof of Proposition 5. Let G and H be graphs satisfying the assumptions of the proposition. Let $\varepsilon > 0$ be such that

$$\alpha\left(1-\varepsilon\left(1+\frac{3}{\kappa}\right)\right)=\alpha'.$$

By Lemma 9, there must exist an edge-coloured graph G_1 such that G_1 is an induced subgraph of G (if we forget about the colouring), and such that $\mathrm{red}(G_1) = H$.

Now, by Proposition 12, there exists M depending only on κ, α and ε, and an induced subgraph H^* of G_1 such that $e(H^*) \geqslant \left(1-\varepsilon\left(1+\frac{3}{\kappa}\right)\right)e(H) \geqslant \alpha' e(G)$ and such that for every $v \in V(H^*)$, we have

$$\deg(v) \leqslant 3M d_{\mathrm{red}}(v), \tag{22}$$

and such that $\mathrm{red}(H^*)$ is a $\frac{\kappa}{3}$-expander.

By (22) and Lemma 10, H^* is a κ'-expander, with $\kappa' = \frac{\kappa}{9M}$, which concludes the proof. □

References

[1] M. Ajtai, J. Komlos, and E. Szemeredi. Deterministic simulation in logspace. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 132–140, New York, NY, USA, 1987. Association for Computing Machinery.

[2] N. Alon. Eigenvalues and expanders. volume 6, pages 83–96. 1986. Theory of computing (Singer Island, Fla., 1984).

[3] D. Chakraborti, J. Kim, J. Kim, M. Kim, and H. Liu. Well-mixing vertices and almost expanders. *Proceedings of the American Mathematical Society*, 150(12):5097–5110, 2022.

[4] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pages 195–199. 1970.

[5] F. R. Chung. *Spectral graph theory.* Number 92. American Mathematical Soc., 1997.

[6] D. Gillman. A Chernoff bound for random walks on expander graphs. *SIAM J. Comput.*, 27(4):1203–1220, 1998.

[7] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561, 2006.

[8] R. Kannan, S. Vempala, and A. Vetta. On clusterings: good, bad and spectral. *J. ACM*, 51(3):497–515, 2004.

[9] M. Krivelevich. Finding and using expanders in locally sparse graphs. *SIAM J. Discrete Math.*, 32(1):611–623, 2018.

[10] M. Krivelevich. Expanders – how to find them, and what to find in them. *Surveys in Combinatorics*, pages 115–142, 2019.

[11] B. Louf. Large expanders in high genus unicellular maps. *Combinatorial Theory* 2(3), 2022.

[12] R. Montgomery. Logarithmically small minors and topological minors. *J. Lond. Math. Soc. (2)*, 91(1):71–88, 2015.

[13] A. Shapira and B. Sudakov. Small complete minors above the extremal edge density. *Combinatorica*, 35(1):75–94, 2015.