Random Structures & Algorithms

WILEY

RESEARCH ARTICLE

Separating Path Systems in Complete Graphs

Cristina G. Fernandes[1] | Guilherme Oliveira Mota[1] | Nicolás Sanhueza-Matamala[2]

[1]Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil | [2]Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile

Correspondence: Nicolás Sanhueza-Matamala (nsanhuezam@udec.cl)

Received: 22 December 2023 | **Revised:** 18 January 2025 | **Accepted:** 25 March 2025

Funding: This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Agencia Nacional de Investigación y Desarrollo, Fundação de Amparo à Pesquisa do Estado de São Paulo, Conselho Nacional de Desenvolvimento Científico e Tecnológico, by a joint project FAPESP and ANID (2019/13364-7), CAPES (Finance Code 001), CNPq (310979/2020-0), CNPq (306620/2020-0, 406248/2021-4, 404315/2023-2), FAPESP (2018/04876-1), and ANID-FONDECYT Iniciación N°11220269 grant and ANID-FONDECYT Regular N°1251121 grant.

Keywords: complete graphs | hypergraphs | separating path systems

ABSTRACT

We prove that in any n-vertex complete graph, there is a collection $\mathcal{P}$ of $(1+o(1))n$ paths that *strongly separates* any pair of distinct edges e, f, meaning that there is a path in $\mathcal{P}$, which contains e but not f. Furthermore, for certain classes of n-vertex αn-regular graphs, we find a collection of $(\sqrt{3\alpha+1}-1+o(1))n$ paths that strongly separates any pair of edges. Both results are best-possible up to the $o(1)$ term.

1 | Introduction

1.1 | Separating Path Systems

Let $\mathcal{P}$ be a family of paths in a graph G. We say that two edges e *and* f are *weakly separated by* $\mathcal{P}$ if there is a path in $\mathcal{P}$, which contains one of these edges but not both. We also say that they are *strongly separated by* $\mathcal{P}$ if there are two paths $P_e, P_f \in \mathcal{P}$ such that P_e contains e but not f, and P_f contains f but not e.

We are interested in the problem of finding "small" families of paths ("path systems") that separate any pair of edges in a given graph. A path system in a graph$G$ is *weak-separating* (resp. *strong-separating*) if all pairs of edges in G are weakly (resp. strongly) separated by it. Let $\mathrm{wsp}(G)$ and $\mathrm{ssp}(G)$, respectively, denote the size of the smallest such families of paths in a graph G. Since every strong-separating path system is also weak-separating, the inequality $\mathrm{wsp}(G) \leq \mathrm{ssp}(G)$ holds for any graph G, but equality is not true in general.

The study of general separating set systems was initiated by Rényi [1] in the 1960s. The variation that considers the separation of edges using subgraphs has been considered many times in the computer science community, motivated by the application of efficiently detecting faulty links in networks [2–5]. The question got renewed interest in the combinatorics community after it was raised by Katona in a conference in 2013 and was considered simultaneously by Falgas–Ravry et al. [6] in its weak version, and by Balogh et al. [7] in its strong version. Both teams conjectured that n-vertex graphs G admit (weak and strong) separating path systems of size linear in n, that is, $\mathrm{wsp}(G), \mathrm{ssp}(G) = O(n)$, and both also observed that an $O(n \log n)$ bound holds. Letzter [8] made substantial progress in this question by showing that all n-vertex graphs G satisfy $\mathrm{ssp}(G) = O(n \log^* n)$. The conjecture was settled by Bonamy et al. [9], who proved that $\mathrm{ssp}(G) \leq 19n$ holds for any n-vertex graph G.

1.2 | Separating Cliques

An interesting open question is to replace the value of '19' in $\mathrm{ssp}(G) \leq 19n$ by the smallest possible number. Perhaps, it could be possible even to replace this value by $1 + o(1)$. Studying separating path systems in complete graphs is particularly relevant since K_n gives the best-known lower bounds for $\mathrm{wsp}(G)$ and $\mathrm{ssp}(G)$ over all n-vertex graphs G (see Section 9 for further discussion). Because of this fact, the behavior of $\mathrm{ssp}(K_n)$ and $\mathrm{wsp}(K_n)$ has been enquired repeatedly by many authors (e.g., [6], Sect 7).

For the weak separation, we know that $\mathrm{wsp}(K_n) \geq n - 1$ (see the remark before Conjecture 1.2 in [6]). For strong separation, mimicking that proof shows that the slightly better bound $\mathrm{ssp}(K_n) \geq n$ holds (this is done in Theorem 2.1). Our first main result shows that this lower bound is asymptotically correct.

Theorem 1.1. *The following holds.*

$$\mathrm{ssp}(K_n) = (1 + o(1))n$$

Let us summarize the history of upper bounds for this problem now. First, we knew that $\mathrm{wsp}(K_n) = O(n)$ [6], Theo 1.3, and then $\mathrm{ssp}(K_n) \leq 2n + 4$ [7], Theo 3. Wickes [10] studied $\mathrm{wsp}(K_n)$ in more detail and showed that $\mathrm{wsp}(K_n) \leq n$ whenever n or $n - 1$ is a prime number and that $\mathrm{wsp}(K_n) \leq (21/16 + o(1))n$ in general. After the preprint version of this work had appeared, Kontogeorgiou and Stein [11] proved that $\mathrm{wsp}(K_n) \leq n + 2$.

The problem of estimating $\mathrm{ssp}(K_n)$ is connected with the older problem of finding *orthogonal double covers* (ODC), which are collections $\mathcal{C}$ of subgraphs of K_n in which every edge appears in exactly two elements of $\mathcal{C}$, and the intersection of any two elements of $\mathcal{C}$ contains exactly one edge. If each graph in $\mathcal{C}$ is isomorphic to some graph H, we say that $\mathcal{C}$ is an *ODC by* H. If H is an n-vertex path and $\mathcal{C}$ is an ODC by H, then each element of $\mathcal{C}$ is a Hamiltonian path, and it is easy to check that $\mathcal{C}$ must contain exactly n paths and forms a strong-separating path system. Moreover, a counting argument (see Remark 2.2) yields that a strong-separating system in K_n with n paths must form an ODC by Hamiltonian paths. Thus, we know that $\mathrm{ssp}(K_n) = n$ if and only if an ODC by Hamiltonian paths exists. This statement is known to be false for $n = 4$ (it can be checked that $\mathrm{ssp}(K_4) = 5$, see [12], Sect 3.4), but is conjectured to be true for all other $n \geq 3$. It is known to be true for infinitely many values of n, in particular, it holds if n can be written as a product of the numbers 5, 9, 13, 17, and 29 [13]. See the survey [12] for more results and details. We discuss this connection further in Section 9.

1.3 | Separating Regular Graphs

Our main result for cliques (Theorem 1.1) follows from a more general result that works for "robustly-connected" graphs, which are *almost regular*, meaning that each vertex has approximately the same number of neighbors. For simplicity, we give the statement for regular graphs here. Let $\alpha \in [0, 1]$ and consider an αn-regular graph G on n vertices. A counting argument (Theorem 2.3) shows that $\mathrm{ssp}(G) \geq (\sqrt{3\alpha + 1} - 1 - o(1))n$ must hold. Our second main result shows that this bound essentially holds with equality if we also assume some vertex-connectivity condition. We say an n-vertex graph G is (δ, L)*-robustly-connected* if, for every $x, y \in V(G)$, there exists $1 \leq \ell \leq L$ such that there are at least δn^{ℓ} (x, y)-paths with exactly ℓ inner vertices each.

Theorem 1.2. *Let $\alpha, \delta \in (0, 1)$ and $L \geq 1$. Suppose that G is an n-vertex graph, which is αn-regular and (δ, L)-robustly-connected. Then*

$$\mathrm{ssp}(G) = (\sqrt{3\alpha + 1} - 1 + o(1))n$$

We note that at least some kind of connectivity is required for a bound like the one in Theorem 1.2. Indeed, the graph G formed by two vertex-disjoint cliques with $n/2$ vertices is $(n/2 - 1)$-regular but has clearly $\mathrm{ssp}(G) = 2 \cdot \mathrm{ssp}(K_{n/2}) \geq n$, whereas Theorem 1.2 would give an incorrect upper bound around $(0.582 + o(1))n$.

Observe that the function $f(\alpha) = \sqrt{3\alpha + 1} - 1$ satisfies $\alpha < f(\alpha) < \sqrt{\alpha} < 1$ for $\alpha \in (0, 1)$, so in particular this shows that all n-vertex graphs G covered by Theorem 1.2 satisfy $\mathrm{ssp}(G) \leq (1 + o(1))n$. From Theorem 1.2, we can obtain as corollaries results for many interesting classes of graphs as balanced complete bipartite graphs, regular graphs with large minimum degree, regular robust expanders, etc. (see Section 8 for details).

1.4 | Outline of the Proof

We summarize the idea behind our proof by focusing on the case of estimating $\mathrm{ssp}(K_n)$. The calculations that give the lower bound $\mathrm{ssp}(K_n) \geq n$ (Theorem 2.1) reveal that, if $\mathrm{ssp}(K_n) = n$ holds, then in an optimal strong-separating path system, each path must be Hamiltonian, each edge must be covered precisely by two paths, and every two different paths intersect precisely on one edge. Guided by this, our approach can be thought conceptually of taking $t = (1 + \varepsilon)n$ different 'labels' and finding an injective assignment $\phi : E(K_n) \to \binom{t}{2}$ where every edge gets two labels. Then, by defining, for each $1 \leq i \leq t$, the subgraph $Q_i \subseteq K_n$ consisting of the edges that received label i, we get that the family $\{Q_i\}_{1 \leq i \leq t}$ will strongly separate the edges of K_n.

To make sure that the graphs Q_i resemble paths, we will obtain the assignment ϕ in a more careful way. We will construct ϕ with the help of an almost perfect matching in an auxiliary hypergraph $\mathcal{H}$. In this case, the hypergraph can be described as follows: We randomly orient the edges of K_n to obtain a digraph D where every vertex gets approximately the same number of incoming and outgoing edges. Then obtain an auxiliary graph B by taking two copies V_1, V_2 of $V(K_n)$ and adding an edge between $u_1 \in V_1$ and $v_2 \in V_2$ if the arc (u, v) appears in D. Next, consider a clique K on a set of vertices $\{1, 2, \dots, n\}$, vertex disjoint from $V_1 \cup V_2$. Form a graph Z by adding every edge between a vertex of $V_1 \cup V_2$ and a vertex in K. Then, if $u_1 \in V_1, v_2 \in V_2, i, j \in V(K)$ and these vertices form a K_4 in Z (say those copies of K_4 are 'valid'), we can interpret that as assigning the edge $uv \in E(K_n)$ to the graphs Q_i and Q_j. Similarly, if we have edge-disjoint valid copies of K_4 in Z, this can be interpreted as assigning different edges of $E(K_n)$ to different pairs Q_i, Q_j; without repeating pairs, and assigning at most two edges adjacent to the same vertex in K_n to the same Q_i. Thus, if we can find edge-disjoint valid copies of K_4, which use all edges between V_1 and V_2, we would have obtained an allocation of all the edges of $E(K_n)$ to pairs of Q_i, Q_j, where each Q_i has maximum degree at most 2. To find such edge-disjoint copies of K_4, we look at the auxiliary 6-graph $\mathcal{H}$ with vertex set $E(Z)$ and each valid copy of K_4 corresponding to an edge in $\mathcal{H}$. By construction, an almost perfect matching in $\mathcal{H}$ will yield graphs $\{Q_i\}_{1\le i\le t}$, which separate "almost all" pairs of edges and have the crucial property that $\Delta(Q_i) \le 2$ for each i. This will ensure that the graphs Q_i are collections of paths and cycles. To find such an almost perfect matching in $\mathcal{H}$, we will use a recent powerful result on hypergraph matchings by Glock et al. [14], which will allow us to gain even more control of the shape of the graphs Q_i by avoiding certain undesirable short cycles.

After this is done, we will have covered and separated most, but not all, of the edges of K_n with the graphs $\{Q_i\}_{1\le i\le t}$, which are collections of paths and cycles. In a next step, we will transform each Q_i by merging (most of) the edges of Q_i into a single path Q_i'. This is done carefully to ensure the path system $\mathcal{Q} = \{Q_i'\}_{1\le i\le t}$ still strongly separates most of the edges of the graph.

In the final step, the subgraph $H \subseteq G$ of edges, which remain unseparated, will be very sparse, and in particular, has very small maximum degree (at most εn). Using a probabilistic argument based on the Lovász Local Lemma, we find a small (of size $O(\varepsilon n)$) strong-separating path system $\mathcal{P}$, which strongly separates H. Then, our final desired path system will be given by $\mathcal{P} \cup \mathcal{Q}$.

In this sketch of the proof, we glossed over some details. In the actual proof (which covers the general case for $G \neq K_n$), the situation is slightly more technical because in an optimal solution, the edges of G need to be covered by 2 or 3 paths (as can be seen from Theorem 2.3). The outline of the proof is the same, but instead, we will use a more intricate auxiliary hypergraph $\mathcal{H}$ (in fact, we use an 8-uniform graph) to find the initial assignment.

1.5 | Organization of the Paper

In Section 2, we give simple counting arguments that yield the lower bounds in Theorems 1.1 and 1.2. Then, we begin the proof of our main result. In Section 3, we gather some probabilistic and hypergraph tools and prove results that will be helpful during the next sections. In Section 4, we find a family of graphs that separates almost all edges of a graph G via a perfect matching in an auxiliary hypergraph. In Section 5, we transform the given families of graphs into paths, keeping some structural properties. In Section 6, we find small path systems that separate the remaining leftover edges. Then, the pieces of the main proofs are put together in Section 6. In Section 8, we describe how to use our main result in some important graph classes, and we finalize with concluding remarks in Section 9.

2 | Lower Bounds

Given a path system $\mathcal{P}$ in a graph G and $e \in E(G)$, let $\mathcal{P}(e) \subseteq \mathcal{P}$ be the paths of $\mathcal{P}$ which contain e. Note that $\mathcal{P}$ is weak-separating if and only if the sets $\mathcal{P}(e)$ are different for all $e \in E(G)$; and $\mathcal{P}$ is strong-separating if and only if no set $\mathcal{P}(e)$ is contained in another $\mathcal{P}(f)$.

Proposition 2.1. *For each $n \ge 3$, $\mathrm{ssp}(K_n) \ge n$.*

Proof. Let $n \ge 3$. Let $\mathcal{P}$ be a strong-separating path system in K_n, and define $\mathcal{P}_1 = \{P \in \mathcal{P} : |E(P)| = 1\}$. Note that

$$\begin{aligned}\sum_{e\in E(K_n)} |\mathcal{P}(e)| &= \sum_{P\in\mathcal{P}} |E(P)| \\ &\le |\mathcal{P}_1| + |\mathcal{P} \setminus \mathcal{P}_1|(n-1) = |\mathcal{P}|(n-1) - |\mathcal{P}_1|(n-2)\end{aligned}$$

where we used that each path can contain at most $n-1$ edges.

Next, let $E_1 \subseteq E(K_n)$ be the set of edges e such that $|\mathcal{P}(e)| = 1$. Note that $|E_1| \le |\mathcal{P}_1|$, because if an edge e is covered by an unique path P, then P cannot cover any other edge f, as otherwise there would be no other path, which covers e and not f, a contradiction to the fact that $\mathcal{P}$ is strong-separating. We have that

$$\sum_{e\in E(K_n)} |\mathcal{P}(e)| \geq |E_1| + 2\left(\binom{n}{2} - |E_1|\right)$$
$$= n(n-1) - |E_1| \geq n(n-1) - |\mathcal{P}_1|$$

which implies that

$$n(n-1) \leq |\mathcal{P}|(n-1) - |\mathcal{P}_1|(n-3) \leq |\mathcal{P}|(n-1)$$

where the last inequality uses $n \geq 3$. This implies that $|\mathcal{P}| \geq n$. □

Remark 2.2. If a strong-separating path system $\mathcal{P}$ in K_n has size n, then all inequalities in the previous proof become equalities. This implies that $0 = |E_1| = |\mathcal{P}_1|$, that every edge of K_n is covered exactly by two different paths, and that every path must be Hamiltonian and intersect every other path exactly once; so $\mathcal{P}$ is an ODC by Hamiltonian paths, as mentioned in the introduction.

Proposition 2.3. *For any $\alpha, \varepsilon \in (0,1]$, the following holds for all sufficiently large n. Let G be an n-vertex graph with $\alpha\binom{n}{2}$ edges. Then*

$$\mathrm{ssp}(G) \geq (\sqrt{3\alpha+1} - 1 - \varepsilon)n$$

Proof. Let α, ε be given, and suppose n is sufficiently large. Given G as in the statement, let $\mathcal{P}$ be a strong-separating path system of size $\mathrm{ssp}(G)$. Suppose β is such that $|\mathcal{P}| = \beta n$ (we know that $\beta \leq 19$ by the result of [9]). We need to show that $\beta \geq \sqrt{3\alpha+1} - 1 - \varepsilon$. Note that

$$\sum_{e\in E(G)} |\mathcal{P}(e)| = \sum_{P\in\mathcal{P}} |E(P)| \leq \beta n(n-1) = 2\beta\binom{n}{2}$$

For $i \in \{1,2\}$, let $E_i \subseteq E(G)$ be the set of edges e such that $|\mathcal{P}(e)| = i$. Then

$$2\beta\binom{n}{2} \geq \sum_{e\in E(G)} |\mathcal{P}(e)| \geq |E_1| + 2|E_2| + 3\left(\alpha\binom{n}{2} - |E_1| - |E_2|\right) = 3\alpha\binom{n}{2} - 2|E_1| - |E_2| \tag{2.1}$$

Since $\mathcal{P}$ is strong-separating, if $e \in E_2$, then the two paths of $\mathcal{P}$ that contain e cannot both contain any other edge $f \in E_2$. Thus, $|E_2| \leq \binom{|\mathcal{P}|}{2} \leq \binom{\beta n}{2} \leq \beta^2\binom{n}{2} + \beta^2 n$. Note that we also have $|E_1| \leq |\mathcal{P}| \leq \beta n$. Applying these bounds on $|E_1|$ and $|E_2|$ in (2.1), we get

$$\beta^2\binom{n}{2} + 2\beta\binom{n}{2} \geq 3\alpha\binom{n}{2} - \beta^2 n - 2\beta n \geq 3\alpha\binom{n}{2} - 400n$$

where in the last step we used $\beta \leq 19$ to get $\beta^2 n + 2\beta n \leq 400n$. Thus, the inequality $\beta^2 + 2\beta \geq 3\alpha - 800/n$ holds. Since $\beta > 0$ and n is sufficiently large, solving this quadratic equation in terms of β gives that $\beta \geq \sqrt{3\alpha+1} - 1 - \varepsilon$, as desired. □

3 | Preliminaries

3.1 | Hypergraph Matchings

We use a recent result by Glock et al. [14] (similar results were obtained also by Delcourt and Postle [15]). This result allows us to find almost perfect matchings in hypergraphs $\mathcal{H}$, which avoid certain "conflicts." Each conflict is a subset of edges $X \subseteq E(\mathcal{H})$, which we do not want to appear together in the matching M, that is, we want $X \nsubseteq M$ for all such conflicts X. We encode these conflicts using an auxiliary "conflict hypergraph" $\mathcal{C}$ whose vertex set is $E(\mathcal{H})$ and each edge is a different conflict, that is, each edge of $\mathcal{C}$ encodes a set of edges of $\mathcal{H}$.

Given a (not necessarily uniform) hypergraph $\mathcal{C}$ and $k \geq 1$, let $\mathcal{C}^{(k)}$ denote the subgraph of $\mathcal{C}$ consisting of all edges of size exactly k. If $\mathcal{C} = \mathcal{C}^{(k)}$, then $\mathcal{C}$ is a *k-graph*. For a hypergraph $\mathcal{H}$ and $j \geq 1$, let $\delta_j(\mathcal{H})$ (resp. $\Delta_j(\mathcal{H})$) be the minimum (resp. maximum) of the number of edges of $\mathcal{H}$, which contain S, taken over all j-sets S of vertices. We say that a hypergraph $\mathcal{H}$ is *$(x \pm y)$-regular* if $x - y \leq \delta_1(\mathcal{H}) \leq \Delta_1(\mathcal{H}) \leq x + y$. Let $N_{\mathcal{H}}(v)$ denote the set of neighbors of v in $\mathcal{H}$. Given a hypergraph $\mathcal{C}$ with $V(\mathcal{C}) = E(\mathcal{H})$, we say $E \subseteq E(\mathcal{H})$ is *$\mathcal{C}$-free* if for every $C \in E(\mathcal{C})$, C is not a subset of E. Also, $\mathcal{C}$ is a *(d, ℓ, ρ)-bounded conflict system* for $\mathcal{H}$ if

C1. $3 \leq |C| \leq \ell$ for each $C \in \mathcal{C}$;

C2. $\Delta_1(\mathcal{C}^{(j)}) \leq \ell d^{j-1}$ for all $3 \leq j \leq \ell$; and

C3. $\Delta_{j'}(\mathcal{C}^{(j)}) \leq \ell d^{j-j'-\rho}$ for all $3 \leq j \leq \ell$ and $2 \leq j' < j$.

We say that a set of edges $Z \subseteq E(\mathcal{H})$ is (d, ρ)-*trackable*[1] if $|Z| \geq d^{1+\rho}$.

Theorem 3.1 ([14], Theo 3.2). *For all $k, \ell \geq 2$, there exists $\rho_0 > 0$ such that for all $\rho \in (0, \rho_0)$, there exists d_0 so that the following holds for all $d \geq d_0$. Suppose $\mathcal{H}$ is a k-graph on $n \leq \exp(d^{\rho^3})$ vertices with $(1 - d^{-\rho})d \leq \delta_1(\mathcal{H}) \leq \Delta_1(\mathcal{H}) \leq d$ and $\Delta_2(\mathcal{H}) \leq d^{1-\rho}$ and suppose C is a (d, ℓ, ρ)-bounded conflict system for $\mathcal{H}$. Suppose $\mathcal{Z}$ is a set of (d, ρ)-trackable sets of edges in $\mathcal{H}$ with $|\mathcal{Z}| \leq \exp(d^{\rho^3})$. Then, there exists a C-free matching $\mathcal{M} \subseteq \mathcal{H}$ of size at least $(1 - d^{-\rho^3})n/k$ with $|Z \cap \mathcal{M}| = (1 \pm d^{-\rho^3})|\mathcal{M}||Z|/|E(\mathcal{H})|$ for all $Z \in \mathcal{Z}$.*

3.2 | Counting Cycles

Let D_n be the complete digraph (having all arcs in both directions). The following lemma is a simple counting argument that will be used later.

Lemma 3.1. *If $R \subseteq E(D_n)$ has $\ell < j$ edges, then there are at most $j^\ell n^{j-\ell-1}$ length-j directed cycles in D_n which contain R.*

Proof. We can assume that R is a proper subgraph of some directed cycle (as otherwise there is nothing to count). Thus, R is a collection of vertex-disjoint paths in D_n with exactly ℓ edges in total. All directed cycles on j vertices, which contain R can be obtained by assigning a number in $\{1, \dots, j\}$ to allocate the starting position for each of the paths in the cycle, and then choosing each of the $j - |V(R)|$ remaining vertices. Note that R can consist of at most ℓ paths, so the first step can be done in at most j^ℓ ways. On the other hand, R spans at least $\ell + 1$ vertices (the minimum number occurs when R is a single path), so there are at most $n^{j-\ell-1}$ ways to choose the vertices in the second step. Therefore, R is contained in at most $j^\ell n^{j-\ell-1}$ length-j directed cycles in D_n. □

3.3 | Probabilistic Tools

In this short section, we state some standard probabilistic tools used in our proof.

Lemma 3.2 (Chernoff's inequalities [16], Remark 2.5, Corollary 2.3 and 2.4). *Let X be a random variable with binomial distribution $B(n, p)$. Let $t \geq 0$. Then,*

i. $\Pr[|X - \mathbf{E}[X]| \geq t] \leq 2\exp(-2t^2/n)$;

ii. *if* $t \leq 3\mathbf{E}[X]/2$, *then* $\Pr[|X - \mathbf{E}[X]| \geq t] \leq 2\exp(-t^2/(3\mathbf{E}[X]))$; *and*

iii. *if* $t \geq 7\mathbf{E}[X]$, *then* $\Pr[X \geq t] \leq \exp(-t)$.

The following concentration inequality will also be useful.

Lemma 3.3 (McDiarmid's inequality [17]). *Let $X_1, \dots, X_M$ be independent random variables, with X_i taking values on a finite set A_i for each $i \in [M]$. Suppose that $f : \prod_{j=1}^M A_j \to \mathbb{R}$ satisfies $|f(x) - f(x')| \leq c_i$ whenever the vectors x and x' differ only in the ith coordinate, for every $i \in [M]$. Consider the random variable $Y = f(X_1, \dots, X_M)$ and $t \geq 0$. Then*

$$\Pr[|Y - \mathbf{E}(Y)| \geq t] \leq 2\exp\left(-\frac{2t^2}{\sum_{j=1}^M c_j^2}\right)$$

3.4 | Building a Base Hypergraph

The next lemma constructs an auxiliary hypergraph which we will use as a base to apply Theorem 3.1 later. To motivate this lemma regarding the proof sketch given in Section 1.4, the hypergraph J we build in the next lemma will play a similar role as the clique K there.

Lemma 3.4. *For any $\alpha, \beta, \lambda > 0$ with $\beta = \sqrt{3\alpha + 1} - 1 < \lambda$, there exists n_0 such that the following holds for every $n \geq n_0$. There exists a 3-graph J such that*

J1. there is a partition $\{U_1, U_2\}$ of $V(J)$ with $|U_1| = \lambda n$ and $|U_2| = \lambda n\beta/2$;

J2. there is a partition $\{J_1, J_2\}$ of $E(J)$ such that
- *$e \subseteq U_1$ for each $e \in J_1$, and*
- *$|e \cap U_1| = 2$ for each $e \in J_2$;*

J3. every pair $\{i, j\} \subseteq U_1$ is contained only in edges of J_1, or in at most one edge of J_2;

J4. $\Delta_2(J) \le \ln^2 n$;

J5. *J has $\alpha\binom{n}{2} \pm n^{2/3}$ edges in total; and*

J6. *J is $(\beta n/\lambda, n^{2/3})$-regular.*

Proof. We begin by observing that $3\alpha = 2\beta + \beta^2$ from the definition of β. Then, defining $d_2 = \beta^2 n/\lambda$ and $d_3 = 3(\alpha - \beta^2)n/(2\lambda)$, we obtain $d_2 + d_3 = \beta n/\lambda$. A $\{2,3\}$-*graph* is a hypergraph whose edges have size either 2 or 3. We say it is an *antichain* if no edge is contained in another. We begin our construction by defining an antichain $\{2,3\}$-graph on a set of size λn. Let U_1 be a set of λn vertices. We claim that there is an antichain $\{2,3\}$-graph I on U_1 such that

F1. each vertex is adjacent to $d_2 \pm n^{2/3}$ edges of size 2 in I;

F2. each vertex is adjacent to $d_3 \pm n^{2/3}$ edges of size 3 in I.

Indeed, define a random graph $I^{(2)}$ on U_1 by including each edge independently with probability $p := \beta^2/\lambda^2 < 1$. Let $\overline{I}$ be the complement of $I^{(2)}$. In expectation, each vertex is contained in around $p|U_1| = d_2$ many edges. A standard application of Chernoff's inequality (Theorem 3.2(i)) shows that, with overwhelmingly large probability, each vertex of U_1 is contained in $d_2 \pm n^{2/3}$ edges of $I^{(2)}$, and thus we can assume that a choice of $I^{(2)}$ is fixed and satisfies that property. Similarly, we can also assume that every vertex is contained in $(1-p)^3\binom{\lambda n}{2} \pm n^{4/3}$ triangles in $\overline{I}$. Next, we form a 3-graph $I^{(3)}$ on U_1 by including each triple of vertices, which forms a triangle in $\overline{I}$ with probability $q := 3(\alpha - \beta^2)/((1-p)^3\lambda^3 n)$. If a vertex x is contained in $d := (1-p)^3\binom{\lambda n}{2}$ triangles in $\overline{I}$, then in expectation it must be contained in $dq = d_3 \pm n^{1/2}$ many 3-edges in $I^{(3)}$. Using Chernoff again (Theorem 3.2(ii)), we can assume that each vertex in $I^{(3)}$ is contained in $d_3 \pm n^{2/3}$ many triangles in $I^{(3)}$. We conclude by taking $I = I^{(2)} \cup I^{(3)}$.

Now, we transform I into a 3-graph. To achieve this, we will add a set U_2 of extra vertices to I and extend each 2-edge of I to a 3-edge by including in it a vertex in U_2. Let U_2 have size $r := \lambda n\beta/2$ and vertices $\{v_1, \dots, v_r\}$. Randomly partition the 2-edges of $I^{(2)}$ into r sets $F_1, \dots, F_r$ by including each edge of $I^{(2)}$ in an F_i with probability $1/r$. Next, define sets of 3-edges $H_1, \dots, H_r$ given by $H_i := \{xyv_i : xy \in F_i\}$.

Let J be the 3-graph on vertex set $U_1 \cup U_2$ whose edges are $I^{(3)} \cup \bigcup_{i=1}^r H_i$. Note that, by construction, J satisfies (J1)–(J3), so it only remains to verify (J4), (J5), and (J6).

We show that (J4) holds. Let x and y be a pair of vertices in $V(J)$ and let us consider the possible cases. If $x, y \in U_1$ and $xy \in I^{(2)}$, then $\deg_J(xy) = 1$ because its only neighbor is v_i (if $xy \in F_i$). If $x, y \in U_1$ and $xy \in \overline{I}$, then $\deg_J(xy)$ is precisely the number of triangles of $I^{(3)}$ that contain xy. This is a random variable with expected value at most $nq = O(1)$. Thus, by Theorem 3.2(iii), $\deg_J(xy) > \ln^2 n$ holds with probability at most $n^{-\ln n}$, so we can comfortably use the union bound to ensure that $\deg_J(xy) \le \ln^2 n$ for every such pair $xy \in \overline{I}$. If $x \in U_1$ and $y \in U_2$, then $y = v_i$ for some $1 \le i \le r$, and $\deg_J(xy)$ is the number of triangles of the form $xzv_i \in H_i$. For a fixed x, there are at most $|U_1| = \lambda n$ choices for z to form an edge $xz \in I^{(2)}$ and recall that each such edge belongs to H_i with probability $1/r = O(1/n)$. Thus, the expected value of $\deg_J(xy)$ is again of the form $O(1)$, and we can conclude the argument in a similar way as before. Finally, if $x, y \in U_2$, then $\deg_J(xy) = 0$ by construction. This finishes the proof of (J4).

Note that $|E(J)| = |E(I^{(2)})| + |E(I^{(3)})|$. From (F1), we deduce that $|E(I^{(2)})|$ is $\lambda n(d_2 \pm n^{2/3})/2 = \beta^2 n^2/2 \pm n^{2/3}$ and, from (F2), we deduce that $|E(I^{(3)})|$ is $\lambda n(d_3 \pm n^{2/3})/3 = (\alpha - \beta^2)n^2/2 \pm n^{2/3}$, so $|E(J)| = \alpha n^2/2 \pm O(n^{2/3})$, which proves (J5).

Now we prove (J6). Let $i \in V(J)$. If $i \in U_1$, then $\deg_J(i) = \deg_I(i)$. Since $d_2 + d_3 = \beta n/\lambda$, we have that $\deg_I(i) = d_2 + d_3 + O(n^{2/3}) = \beta n/\lambda + O(n^{2/3})$. Assume now that $i \in U_2$. Recall that we defined J in a way that each vertex of U_2 belongs to $|E(I^{(2)})|/r$ edges, and we have

$$\frac{|E(I^{(2)})|}{r} = \frac{\beta^2\binom{n}{2} \pm O(n^{4/3})}{\lambda n\beta/2} = \frac{\beta n}{\lambda} \pm O(n^{1/3})$$

which concludes the proof of the lemma. □

4 | Separating Almost All Edges

In this section, we show how to separate most pairs of edges of robustly connected graphs by paths and cycles, guaranteeing additional structural properties.

In what follows, let $\varepsilon, \delta > 0$, let L be an integer, and let G be an n-vertex graph. A *2-matching* in G is a collection of vertex-disjoint cycles and paths in G. We say a 2-matching Q in G is *(δ, L)-robustly-connected* if, for every $x, y \in V(Q)$, there exists ℓ with $1 \leq \ell \leq L$ such that there are at least δn^{ℓ} (x, y)-paths with exactly ℓ inner vertices each, all in $V(G) \setminus V(Q)$. Furthermore, a collection $\mathcal{Q}$ of 2-matchings in G is *(δ, L)-robustly-connected* if each $Q \in \mathcal{Q}$ is (δ, L)-robustly-connected. A 2-matching Q in G is *ε-compact* if each cycle in Q has length at least $1/\varepsilon$ and Q contains at most εn paths. For a collection $\mathcal{Q}$ of 2-matchings in G, we say $\mathcal{Q}$ is *ε-compact* if each Q in $\mathcal{Q}$ is ε-compact.

Let $\mathcal{Q} = \{Q_1, \dots, Q_t\}$ be a collection of subgraphs of G. We use $E(\mathcal{Q})$ to denote the set $\bigcup_{i=1}^{t} E(Q_i)$. We say $\mathcal{Q}$ *separates an edge e from all other edges of G* if the set $\{i : e \in E(Q_i)\}$ is not contained in the set $\{j : f \in E(Q_j)\}$ for each $f \in E(G) \setminus \{e\}$. Clearly, if an edge e is separated from all other edges of G by $\mathcal{Q}$, then $e \in E(\mathcal{Q})$. We also say that $\mathcal{Q}$ *strongly separates* a set E' of edges if, for every distinct $e, f \in E'$, the sets $\{i : e \in E(Q_i)\}$ and $\{j : f \in E(Q_j)\}$ are not contained in each other.

For brevity, we put together some of the above definitions in one single concept that will be used in the next result and also in some lemmas in Section 5 (see Lemmas 5.1 and 5.2).

Definition 4.1. Given a graph G and δ, L, β, and ε, we say a collection of 2-matchings $\mathcal{Q}$ is a $(\delta, L, \beta, \varepsilon)$-separator for G if the following holds.

Q1. $\mathcal{Q}$ is ε-compact and (δ, L)-robustly-connected;

Q2. $|\mathcal{Q}| = \beta n$ and $\mathcal{Q}$ strongly separates $E(\mathcal{Q})$;

Q3. each vertex in G is the endpoint of at most εn paths among all $Q \in \mathcal{Q}$;

Q4. each $e \in E(\mathcal{Q})$ is contained in at most three of the 2-matchings in $\mathcal{Q}$; and

Q5. $\Delta(G - E(\mathcal{Q})) \leq \varepsilon n$.

In the next result, we show that large enough (δ, L)-robustly-connected "almost regular" graphs contain a suitable collection of 2-matchings that is a $(\varepsilon^{\ell}\delta/2, L, \beta/(1-\varepsilon), \varepsilon')$-separator, for any ε and ε'.

Lemma 4.2. *Let $1/n \ll \varepsilon, \varepsilon', \alpha, \delta, 1/L, \rho$. Let $\beta = \sqrt{3\alpha + 1} - 1$. If G is an n-vertex $(\alpha n \pm n^{1-\rho})$-regular graph that is (δ, L)-robustly-connected, then there exists a $(\varepsilon^{\ell}\delta/2, L, \beta/(1-\varepsilon), \varepsilon')$-separator for G.*

Proof. Our proof has five steps. First, we build an auxiliary hypergraph $\mathcal{H}$ such that a large matching $M \subseteq \mathcal{H}$, which avoids certain conflicts, yields a family of subgraphs of G with the desired properties. We wish to apply Theorem 3.1 to find such a matching. In the second step, we verify that $\mathcal{H}$ satisfies the hypotheses of Theorem 3.1. In the third step, we define our conflict hypergraph $\mathcal{C}$. In the fourth step, we define some test sets and prove they are trackable. Having done this, we are ready to apply Theorem 3.1, which is done in the last step. Then, we verify that the construction gives the desired graphs. From now on, we can assume that ρ is sufficiently small (since that only weakens our assumptions). Also, we assume that n is sufficiently large with respect to $\varepsilon, \varepsilon', \alpha, \delta, L, \rho$ so that every calculation that requires it is valid.

Step 1: Constructing the auxiliary hypergraph. Obtain an oriented graph D by orienting each edge of G uniformly at random. Each vertex v has expected in-degree and out-degree $d_G(v)/2 = (\alpha n \pm n^{1-\rho})/2$. So, by Chernoff's inequality (Theorem 3.2(i)) and a union bound, we can assume that in D every vertex has in-degree and out-degree of the form $\alpha n/2 \pm 2n^{1-\rho}$.

Next, consider an auxiliary bipartite graph B whose clusters are copies V_1 and V_2 of $V(G)$, where each vertex $x \in V(G)$ is represented by two copies $x_1 \in V_1$ and $x_2 \in V_2$, and such that $x_1y_2 \in E(B)$ if and only if $(x, y) \in E(D)$. Thus, we have that $|E(B)| = |E(G)| = \alpha\binom{n}{2} \pm n^{2-\rho}$, because G is $(\alpha n \pm n^{1-\rho})$-regular. Finally, let $\lambda = \beta/(1-\varepsilon)$. Apply Theorem 3.4 with α, β, γ to obtain a 3-graph J that satisfies (J1)–(J6) and assume that $U_1 = [\lambda n]$ and $V(J) = [|V(J)|]$.

Now, we build an initial auxiliary 8-graph $\mathcal{H}'$ as follows. Let Z be the complete bipartite graph between clusters $V(B)$ and $V(J)$. The vertex set of $\mathcal{H}'$ is $E(B) \cup E(J) \cup E(Z)$. Each edge in $\mathcal{H}'$ is determined by a choice $x_1y_2 \in E(B)$ and $ijk \in E(J)$, which form an edge together with the 6 edges in Z that join x_1 and y_2 to i, j, and k. More precisely, the edge determined by $x_1y_2 \in E(B)$ and $ijk \in E(J)$ is $\Phi(x_1y_2, ijk) := \{x_1y_2, ijk, x_1i, x_1j, x_1k, y_2i, y_2j, y_2k\}$; and the edge set of $\mathcal{H}'$ is given by

$$E(\mathcal{H}') = \{\Phi(x_1y_2, ijk) : x_1y_2 \in E(B), ijk \in E(J)\}$$

The idea behind the construction of $\mathcal{H}'$ is as follows: Suppose that M is a matching in $\mathcal{H}'$, and that $x_1y_2 \in E(B)$ is covered by M and appears together with $ijk \in E(J)$ in an edge of M. By (J2), $\{i, j, k\} \cap U_1$ has size 2 or 3. Recall that we want to obtain a collection $\mathcal{Q} := \{Q_1, \dots, Q_t\}$ of 2-matchings in G, where $t = \lambda n$, satisfying (Q1)–(Q5). We will add edges $xy \in E(G)$ such that $\Phi(x_1y_2, ijk) \in M$ or $\Phi(y_1x_2, ijk) \in M$ to the graphs Q_a if $a \in \{i, j, k\} \cap U_1$.

By construction, and since M is a matching, at most one edge in B involving x_1 (resp. x_2) appears in an edge of M together with some $a \in U_1$. By considering the contributions of the two copies $x_1, x_2 \in V(B)$ of a vertex $x \in V(G)$, this means that the subgraphs $Q_a \subseteq G$ have maximum degree 2, and thus these graphs are 2-matchings in G, as we wanted. By construction and property (J2), each edge in $E(G)$ belongs to either 0, 2, or 3 graphs Q_a. Importantly, property (J3) implies that, for two distinct edges $e, f \in E(G)$, no two nonempty sets of the type $\{a : e \in E(Q_a)\}$ and $\{b : f \in E(Q_b)\}$ can be contained in each other. Straightforward calculations reveal the degrees of the vertices in $V(\mathcal{H}')$.

Claim 4.3. The following hold.

(i) $\deg_{\mathcal{H}'}(x_1y_2) = \alpha\binom{n}{2} \pm n^{2/3}$ for each $x_1y_2 \in E(B)$;

(ii) $\deg_{\mathcal{H}'}(ijk) = \alpha\binom{n}{2} \pm n^{2-\rho}$ for each $ijk \in E(J)$; and

(iii) $\deg_{\mathcal{H}'}(x_a i) = \frac{\alpha\beta}{\lambda}\binom{n}{2} \pm 2n^{2-\rho}$ for each $x_a i \in E(Z)$. □

Proof of the Claim. The first two points can be easily verified: given any $x_1y_2 \in E(B)$, by construction $d_{\mathcal{H}'}(x_1y_2)$ is the number of edges in J, which is $\alpha\binom{n}{2} \pm n^{2/3}$ by (J5). Moreover, given any $ijk \in E(J)$, $d_{\mathcal{H}'}(ijk)$ is the number of edges of $E(B)$, which is $\alpha\binom{n}{2} \pm n^{2-\rho}$ by construction.

Finally, consider $x_a \in V_1$ (the case $x_a \in V_2$ is symmetric) and $i \in V(J)$. The degree $\deg_{\mathcal{H}'}(x_a i)$ corresponds to the edges $\Phi(x_a y_2, ijk)$ with $y_2 \in N_B(x_1)$ and $jk \in N_J(i)$. Next, we estimate the number of valid choices for y_2 and jk. There are $\deg_B(x_1) = d^+_D(x) = \alpha n/2 \pm 2n^{1-\rho}$ possible choices for y_2, and there are $\deg_J(i) = \beta n/\lambda \pm n^{2/3}$ possible choices for jk by (J6). Thus, we deduce that

$$\deg_{\mathcal{H}'}(x_1 i) = \left(\frac{\alpha n}{2} \pm 2n^{1-\rho}\right)\left(\frac{\beta n}{\lambda} \pm n^{2/3}\right) = \frac{\alpha\beta}{\lambda}\binom{n}{2} \pm 2n^{2-\rho}$$

as desired. □

Since $\mathcal{H}'$ is not quite regular, we will actually work with a carefully chosen subgraph $\mathcal{H}$ of $\mathcal{H}'$. Let $p := \beta/\lambda = 1 - \varepsilon$. For each $i \in V(J)$, select a subset $X_i \subseteq V(G)$ by including in X_i each vertex of G independently at random with probability p. This defines a family $\{X_i : i \in V(J)\}$ of subsets of $V(G)$. For each $x \in V(G)$, consider the random set $Y_x = \{i \in V(J) : x \in X_i\}$. Finally, let $\mathcal{H} \subseteq \mathcal{H}'$ be the induced subgraph of $\mathcal{H}'$ obtained after removing all vertices $x_1 i, x_2 i \in E(Z)$ whenever $x \notin X_i$ (or, equivalently, $i \notin Y_x$). Thus, we have that

$$E(\mathcal{H}) = \left\{\Phi(x_1y_2, ijk) : x_1y_2 \in E(B), ijk \in E(J), \{x, y\} \subseteq X_i \cap X_j \cap X_k\right\}$$

Claim 4.4. The following hold simultaneously with positive probability.

(i) X_i has $pn \pm n^{2/3}$ vertices of G for each $i \in V(J)$;

(ii) Y_x has $3\alpha n/2 \pm n^{2/3}$ vertices of J for each $x \in V(G)$;

(iii) $\mathcal{H}$ is $(p^6\alpha\binom{n}{2} \pm 2n^{2-\rho})$-regular; and

(iv) for each $i \in V(J)$ and each pair of distinct vertices $x, y \in X_i$, there exists ℓ with $1 \leq \ell \leq L$ such that there are at least $\varepsilon^\ell \delta n^\ell/2$ (x, y)-paths in G with exactly ℓ inner vertices each, all in $V(G) \setminus X_i$. □

Proof of the Claim. Item (i) follows directly from Chernoff's inequality (Theorem 3.2) as, for each $i \in V(J)$, we have $\mathbf{E}[|X_i|] = pn$. For (ii), note that $|V(J)| = \lambda(1 + \beta/2)n$ by (J1) and, for any $x \in V(G)$, we have

$$\mathbf{E}[|Y_x|] = p|V(J)| = \left(\beta + \frac{\beta^2}{2}\right)n = \frac{3\alpha n}{2}$$

Then (ii) also follows from Chernoff's inequality. For (iii), observe that any given edge $\Phi(x_1y_2, ijk) \in \mathcal{H}'$ survives in $\mathcal{H}$ with probability p^6. Using this, we easily see that $\mathbf{E}[\deg_{\mathcal{H}}(x_1y_2)] = p^6 \deg_{\mathcal{H}'}(x_1y_2) = p^6\alpha\binom{n}{2} \pm n^{2-\rho}$ for each $x_1y_2 \in E(B)$; a similar calculation holds for $\mathbf{E}[\deg_{\mathcal{H}}(ijk)]$ for any $ijk \in E(J)$. It remains to calculate the expected degree of the edges in $E(Z) \cap V(\mathcal{H})$. Let $x_a i \in E(Z)$. Conditioning on the event that $x \in X_i$ (and thus that $x_a i \in V(\mathcal{H})$), each edge in $\mathcal{H}'$ containing $x_a i$ survives with probability p^5. Using this, we obtain that

$$\mathbf{E}[\deg_{\mathcal{H}}(x_a i)] = p^5 \deg_{\mathcal{H}'}(x_a i) = p^5\left(\frac{\alpha\beta}{\lambda}\binom{n}{2} \pm 2n^{2-\rho}\right)$$
$$= p^6\alpha\binom{n}{2} \pm 2n^{2-\rho}$$

So (iii) follows from Chernoff's inequality.

In the remainder of the proof, we use McDiarmid's inequality (Theorem 3.3) to check that (iv) holds. Given $i \in V(J)$, since G is (δ, L)-robustly-connected, for each pair of distinct vertices $x, y \in X_i$, there exists ℓ with $1 \le \ell \le L$ such that there are at least δn^ℓ (x, y)-paths in G with exactly ℓ inner vertices each. Hence, the expected number of such paths with all internal vertices in $V(G) \setminus X_i$ is at least $(1-p)^\ell \delta n^\ell$. Since the removal or addition of a vertex in X_i changes the number of (x, y)-paths by $O(n^{\ell-1})$ paths, one can check by using McDiarmid's inequality that, for a given $i \in V(J)$ and a pair $x, y \in X_i$, the probability that we have less than $\varepsilon^\ell \delta n^\ell/2$ (x, y)-paths with exactly ℓ inner vertices, all in $V(G) \setminus X_i$, is $\exp(-\Omega(n))$. Since $|V(J)| = O(n)$ and there are $O(n^2)$ possible pairs $x, y \in X_i$, item (iv) follows from the union bound. □

From now on, we assume that the sets $\{X_i : i \in V(J)\}$, $\{Y_x : x \in V(G)\}$, and the hypergraph $\mathcal{H}$ satisfy properties (i)–(iv) of Claim 4.4.

Step 2: Verifying properties of $\mathcal{H}$. We start by defining $d := \Delta_1(\mathcal{H})$. Note that from Claim 4.4(iii), we have $d = p^6\alpha\binom{n}{2} \pm 2n^{2-\rho}$. We will apply Theorem 3.1 to $\mathcal{H}$. The following claim guarantees that $\mathcal{H}$ satisfies the required hypotheses.

Claim 4.5. The following facts about $\mathcal{H}$ hold.

- (H1) $\mathcal{H}$ has at most $\exp(d^{\rho^3})$ vertices;
- (H2) $d(1 - d^{-\rho/3}) \le \delta_1(\mathcal{H}) \le \Delta_1(\mathcal{H}) = d$; and
- (H3) $\Delta_2(\mathcal{H}) \le d^{2/3}$. □

Proof of the Claim. Item (H1) follows from the fact that $|V(\mathcal{H})| = |E(B)| + |E(J)| + |E(Z)| \le \alpha n^2 + \alpha n^2 + 2n(\lambda + \lambda\beta/2)n \le \exp(d^{\rho^3})$, where the last inequality holds with a lot of room to spare.

For (H2), the upper bound follows from the definition of d and the lower bound follows from $\delta_1(\mathcal{H}) \ge p^6\alpha\binom{n}{2} - 2n^{2-\rho} \ge d - d^{1-\rho/3}$.

It remains to verify that $\Delta_2(\mathcal{H}) \le d^{2/3}$. This will require some work. First, note that each edge in $\mathcal{H}$ is of the form $\Phi(x_1y_2, ijk)$ for $x_1 \in V_1, y_2 \in V_2$, and $ijk \in V(J)$. We need to select two distinct vertices e, f in $V(\mathcal{H})$ and calculate $\deg_{\mathcal{H}}(e, f)$. A vertex e of $\mathcal{H}$ can belong to $E(B)$, $E(J)$, or $E(Z)$. We consider all the six possible combinations for e, f.

Let $e, f \in V(\mathcal{H})$. Since each edge of $\mathcal{H}$ is of type $\Phi(x_1y_2, ijk)$ for $x_1y_2 \in E(B)$ and $ijk \in E(J)$, and each of these contains exactly one vertex in $E(B)$ and one vertex in $E(J)$, for $e, f \in E(B)$ or $e, f \in E(J)$, we have $\deg_{\mathcal{H}}(e, f) = 0$. Furthermore, since $x_1y_2 \in E(B)$ and $ijk \in E(J)$ completely determine the edge $\Phi(x_1y_2, ijk)$, if $e \in E(B)$ and $f \in E(J)$, then we have $\deg_{\mathcal{H}}(e, f) \le 1$.

In view of the above discussion, we may assume that $e \in E(Z)$, and without loss of generality, we assume $e = x_1 i$ for some $x_1 \in V_1$, and $i \in V(J)$. There are now three cases to consider, depending whether f belongs to $E(B)$, $E(J)$, or $E(Z)$.

Suppose first that $f = x_1y_2 \in E(B)$. We will count the number of pairs $\{j, k\}$ such that $\Phi(x_1y_2, ijk)$ is an edge of $\mathcal{H}$. In particular, it must happen that $jk \in N_J(i)$, thus $\deg_{\mathcal{H}}(e, f) \le \deg_J(i) \le n \le d^{2/3}$. Similarly, if $f = ijk \in E(J)$, then we count the number of vertices $y_2 \in V_2$ such that $\Phi(x_1y_2, ijk)$ is an edge of $\mathcal{H}$, which is at most $\deg_B(x_1) \le n \le d^{2/3}$.

Finally, suppose that $f \in E(Z)$ and recall that $e = x_1 i$. If $f = y_2 j$ with $y_2 \in V_2$ and $j \in V(J)$, then $\deg_H(e, f)$ is the number of edges of J containing i and j. In the worst case, $i = j$, we have $\deg_H(e, f) = \deg_J(i) \le n \le d^{2/3}$. On the other hand, if $f = x_1 j$ with $j \in V(J)$, we have $j \ne i$ and then to estimate $\deg_H(e, f)$ we need to count the number of $y_2 \in N_B(x_1) \subseteq V_2$ and the number of $k \in N_J(ij)$. The number of choices for y_2 is at most n and the number of choices for k is at most $\Delta_2(J) \le \log^2 n$. Therefore, $\deg_H(e, f) \le n\log^2 n \le d^{2/3}$, as required. □

Step 3: Setting the conflicts. We must ensure that the collection $\mathcal{Q}$ of 2-matchings we want to obtain is ε'-compact: each 2-matching in $\mathcal{Q}$ has at most $\varepsilon' n$ paths and each cycle in $\mathcal{Q}$ has length at least $1/\varepsilon'$. This condition on the cycle lengths will be encoded by using conflicts.

Recall that D is the oriented graph obtained by orienting each edge of G uniformly at random. In what follows, let $r := 1/\varepsilon'$. We define our conflict hypergraph C on vertex set $E(\mathcal{H})$ and edge set defined as follows: For each ℓ with $3 \le \ell \le r$, each ℓ-length directed cycle $C \subseteq D$ with vertices $\{v^1, \dots, v^\ell\}$, each $i \in U_1$, and each $j_1k_1, \dots, j_\ell k_\ell \in N_J(i)$, we define the following edge:

$$F(C, i, j_1k_1, \dots, j_\ell k_\ell) = \{\Phi(v_1^a v_2^{a+1}, i\ j_a k_a) : 1 \le a \le \ell\}$$

where $v_2^{\ell+1} = v_2^1$. Note that $F(C, i, j_1k_1, \dots, j_\ell k_\ell)$ corresponds to a set of ℓ edges of $\mathcal{H}'$, associated to the triples $(i\ j_1k_1), \dots, (i\ j_\ell k_\ell)$ and the edges of the ℓ-length directed cycle C in D. In such a case, we say i is the *monochromatic color* of the conflicting cycle C. The edges of the conflict hypergraph $\mathcal{C}$ consist of all edges of type $F(C, i, j_1k_1, \dots, j_\ell k_\ell)$ which are contained in $E(\mathcal{H})$. The next claim establishes that $\mathcal{C}$ is a (d, r, ρ)-bounded conflict system for $\mathcal{H}$.

Claim 4.6. The following facts about $\mathcal{C}$ hold.

(C1) $3 \le |F| \le r$ for each $F \in \mathcal{C}$; and

(C2) $\Delta_{j'}(\mathcal{C}^{(j)}) \le rd^{j-j'-\rho}$ for each $3 \le j \le r$ and $1 \le j' < j$. □

Proof of the Claim. Fact (C1) is immediate from the construction of $\mathcal{C}$.

In order to prove (C2), fix j and j' with $3 \le j \le r$ and $1 \le j' < j$. To prove that $\Delta_{j'}(\mathcal{C}^{(j)}) \le rd^{j-j'-\rho}$, we need to show that any set of j' edges of $\mathcal{H}$ is contained in at most $rd^{j-j'-\rho}$ conflicts of size j in $\mathcal{C}$. Let $\mathcal{R}$ be any set of j' edges in $\mathcal{H}$, say $\mathcal{R} = \{\Phi(x_1^a y_2^a, i^a j^a k^a) : 1 \le a \le j'\}$. We want to bound the degree of $\mathcal{R}$ in $\mathcal{C}^{(j)}$. Each conflict of size j is defined by a length-j directed cycle in D, the monochromatic color of the conflict, and a corresponding choice of labels for each edge in the cycle; now we estimate the number of valid choices for each of these three elements.

We begin by estimating how many possibilities there are for choosing a suitable cycle. Note that $R := \{(x^a, y^a) : 1 \le a \le j'\}$ is a set of at most j' edges of D. If $|R| < j'$, then there are repeated edges from D in R, and in this case the degree of $\mathcal{R}$ in $\mathcal{C}^{(j)}$ is zero. So we can assume that $|R| = j'$ and, by Theorem 3.1, there are at most $j^{j'} n^{j-j'-1}$ length-j directed cycles in D which contain R.

Now, we consider the possible choices for the monochromatic color i of the conflicting cycle. Note that if there is no common $i \in V(J)$ among all labels $i^a j^a k^a$ for $1 \le a \le j'$, the degree of $\mathcal{R}$ in $\mathcal{C}^{(j)}$ is zero, because there is no available "monochromatic color" at all. This also implies that there are at most three possible choices for i because it must be one of the three labels which belong to $i^1 j^1 k^1$, say.

Having fixed a directed cycle C, which contains R, and a monochromatic color i for the conflict, now we count the number of labels associated with each edge of C. For edges of $\mathcal{R}$, the choices are already given, and for the remaining $j - j'$ edges of C not in $\mathcal{R}$, the labels must be chosen among the neighbors of i in the hypergraph J. Since i has $\beta n/\lambda \pm n^{2/3} \le n$ neighbors in J by (J6), in this step we have at most $n^{j-j'}$ possible choices. Therefore,

$$\Delta_{j'}(\mathcal{C}^{(j)}) \le j^{j'} n^{j-j'-1} \cdot 3 \cdot n^{j-j'} = 3j^{j'} n^{2(j-j')-1} \le rd^{j-j'-\rho}$$

where in the last step we used that $d = \Theta(n^2)$ and n is sufficiently large. □

Step 4: Setting the test sets. For each $x_1 \in V_1 \subseteq V(B)$ and each $y_2 \in V_2 \subseteq V(B)$, define $Z_{x_1} = \{\Phi(x_1y_2, ijk) \in E(\mathcal{H}) : y_2 \in N_B(x_1), ijk \in E(J)\}$ and define Z_{y_2} in a similar manner. Furthermore, define $Z_i = \{\Phi(x_1y_2, ijk) \in E(\mathcal{H}) : x, y \in X_i\}$ for each $i \in V(J)$. We claim that $\mathcal{Z} := \{Z_{x_1} : x_1 \in V_1\} \cup \{Z_{y_2} : y_2 \in V_2\} \cup \{Z_i : i \in V(J)\}$ is a suitable family of trackable sets. Specifically, the next claim shows that $\mathcal{Z}$ is not very large and has only (d, ρ)-trackable sets.

Claim 4.7. The following facts about $\mathcal{Z}$ hold.

(Z1) $|\mathcal{Z}| \le \exp(d^{\rho^3})$; and

(Z2) each $Z \in \mathcal{Z}$ is (d, ρ)-trackable. □

Proof of the Claim. Because $|\mathcal{Z}| = |V(B)| + |V(J)| \le 3n$ and $d = \Theta(n^2)$, we have $|\mathcal{Z}| \le \exp(d^{\rho^3})$. It remains to check (Z2), which means to prove that $|Z_v| \ge d^{1+\rho}$ for each $v \in V_1 \cup V_2$, and $|Z_i| \ge d^{1+\rho}$ for each $i \in V(J)$.

First, suppose $v = x_1 \in V_1$. From (H2) and $\delta(B) \ge \alpha n/2 - 2n^{1-\rho} \ge \alpha n/3$, we have

$$\begin{aligned} |Z_{x_1}| &= \sum_{y_2 \in N_B(x_1)} \deg_{\mathcal{H}}(x_1y_2) \\ &\ge \delta_1(\mathcal{H})|N_B(x_1)| \ge d(1 - d^{-\rho/3})\alpha n/3 \ge d^{1+\rho} \end{aligned}$$

where in the last step we used that $d = \Theta(n^2)$ and that n is large. The calculations for $v = y_2 \in V_2$ are identical. Next, we note that for any $i \in V(J)$ we have

$$|Z_i| = \sum_{x \in X_i} \deg_{\mathcal{H}}(x_1 i) \geq \delta_1(\mathcal{H})|X_i|$$
$$\geq d(1 - d^{-\rho/3})(pn - n^{2/3}) \geq d^{1+\rho}$$

□

Step 5: Finishing the proof. Recall that $d = \Delta_1(\mathcal{H})$. By Claims 4.5, 4.6, and 4.7, we can apply Theorem 3.1 to $\mathcal{H}$, using $\mathcal{C}$ as a conflict system and $\mathcal{Z}$ as a set of trackable sets, and $\rho/3$ in place of ρ. By doing so, we obtain a matching $\mathcal{M} \subseteq \mathcal{H}$ such that

i. $\mathcal{M}$ is $\mathcal{C}$-free,

ii. $\mathcal{M}$ has size at least $(1 - d^{-(\rho/3)^3})|V(\mathcal{H})|/8$, and

iii. $|Z_a \cap \mathcal{M}| = (1 \pm d^{-(\rho/3)^3})|\mathcal{M}||Z_a|/|E(\mathcal{H})|$ for each $a \in V(B) \cup V(J)$.

Recall that $\lambda = \beta/(1-\varepsilon)$ and let $t = \lambda n$. Using $\mathcal{M}$, we define the graphs $\{Q_i\}_{i=1}^t$ as follows. For an edge $x_1y_2 \in E(B)$, suppose there exists $ijk \in E(J)$ such that $\Phi(x_1y_2, ijk) \in \mathcal{M}$. In that case, we will add the edge $xy \in E(G)$ to the graph Q_a such that $a \in \{i,j,k\} \cap U_1$. To finish, we verify that $\mathcal{Q} = \{Q_i\}_{i=1}^t$ is an $(\varepsilon^\ell \delta/2, L, \lambda, \varepsilon')$-separator for G, which means we need to show that $\mathcal{Q}$ is a collection of 2-matchings in G that satisfies (Q1)–(Q5) with $\varepsilon^\ell \delta/2$, λ and ε' in the place of δ, β and ε, respectively.

We start by verifying that $\mathcal{Q}$ is a collection of 2-matchings. Note that, for each $1 \leq i \leq t$, the graph Q_i has maximum degree at most 2. Indeed, let $x \in V(G)$ be any vertex. Since $\mathcal{M}$ is a matching in $\mathcal{H}$, at most two edges in $\mathcal{M}$ can cover the vertices $x_1i, x_2i \in V(\mathcal{H})$; and this will yield at most two edges adjacent to x belonging to Q_i.

Now we verify that (Q1) holds. First, we check that $\mathcal{Q}$ is ε'-compact, that is, each 2-matching in $\mathcal{Q}$ has at most $\varepsilon' n$ paths and each cycle in $\mathcal{Q}$ has length at least $1/\varepsilon'$. The latter holds because we avoided the conflicts in $\mathcal{C}$. More precisely, an ℓ-cycle in Q_i corresponds to a sequence of ℓ edges, all of which are in Q_i. This means the cycle was formed from a length-ℓ directed cycle in D, all of whose edges were joined (via $\mathcal{M}$) to triples in J, all containing vertex $i \in V(J)$. Recall that $r = 1/\varepsilon'$. If $\ell \leq r$, this forms a conflict in $\mathcal{C}$, so, as $\mathcal{M}$ is $\mathcal{C}$-free, we deduce that $\ell > r$. To check that Q_i has few paths, first observe that $V(Q_i) \subseteq X_i$ for each $i \in V(J)$. Indeed, if $xy \in E(Q_i)$, then we have that, say, $(x,y) \in E(D)$ and $\Phi(x_1y_2, ijk)$ is an edge in $\mathcal{M}$ for some j,k. But, since $\mathcal{M} \subseteq \mathcal{H}$, by the definition of $\mathcal{H}$, we have $x, y \subseteq X_i$, as required. Now, from (ii), (iii), the fact that $\mathcal{H}$ is an 8-graph close to d-regular, and $|Z_i| \leq d|X_i|$, we have that $|E(Q_i)| = |Z_i \cap \mathcal{M}| \geq (1 - \varepsilon'/2)|X_i| \geq (1 - \varepsilon'/2)|V(Q_i)|$, and then the number of degree-one vertices in Q_i is at most $2(|V(Q_i)| - |E(Q_i)|) \leq \varepsilon'|V(Q_i)| \leq \varepsilon' n$. To see the second part of (Q1), we need to show that $\mathcal{Q}$ is $(\varepsilon^\ell \delta/2, L)$-robustly-connected. Because $V(Q_i) \subseteq X_i$, we deduce from Claim 4.4(iv) that Q_i is $(\varepsilon^\ell \delta/2, L)$-robustly-connected, as required. We conclude that (Q1) holds.

We have already stated that $|\mathcal{Q}| = \lambda n$, so the first part of (Q2) holds. The second part of (Q2) can be checked as follows: Let e, f be distinct edges of $E(\mathcal{Q})$. Thus, there are orientations $(x,y), (x',y') \in E(D)$ of e, f respectively, and edges $ijk, i'j'k' \in E(J)$ such that $\Phi(x_1y_2, ijk)$ and $\Phi(x'_1y'_2, i'j'k')$ belong to $\mathcal{M}$. We have, respectively, that $A_e := \{a : e \in E(Q_a)\} = \{i,j,k\} \cap U_1$ and $A_f := \{a : f \in E(Q_a)\} = \{i',j',k'\} \cap U_1$. For a contradiction, suppose $A_e \subseteq A_f$. If $|A_e| = 3$, then we would have that $ijk = i'j'k'$, contradicting that $\mathcal{M}$ is a matching, so $|A_e| = 2$; say, $A_e = \{i,j\}$, and ij is a pair in V_1 (from the construction of J). We recall that by (J3) no pair ij is contained both in an edge with two vertices in V_1 and at the same time in an edge with with three vertices in V_1, so this rules out the case $|A_f| = 3$. Thus, we can only have $A_e = A_f = \{i,j\}$. But again (J3) implies ij is contained in a unique edge in J, say, ijr. This implies that $ijk = i'j'k' = ijr$, contradicting the fact that $\mathcal{M}$ is a matching. Therefore $\mathcal{Q}$ strongly separates $E(\mathcal{Q})$, and (Q2) holds.

To prove (Q3), let $x \in V(G)$. Recall that $Y_x \subseteq V(J)$ is the random set $Y_x = \{i : x \in X_i\}$ and from Claim 4.4(ii) we have that $|Y_x| = 3\alpha n/2 \pm n^{2/3}$. Note that if x is the end of a path in some 2-matching Q_i, then there is an edge $\Phi(x_1y_2, ijk)$ in $Z_{x_1} \cap \mathcal{M}$, but no edge $\Phi(x_2y_1, ijk)$ is in $Z_{x_2} \cap \mathcal{M}$; or there is an edge $\Phi(x_2y_1, ijk)$ in $Z_{x_2} \cap \mathcal{M}$, but no edge $\Phi(x_1y_2, ijk)$ is in $Z_{x_1} \cap \mathcal{M}$. This motivates the following definition: for each $x \in V(G)$, a set $F(x_1)$ of indexes $i \in V(J)$ such that there is an edge $\Phi(x_1y_2, ijk)$ in $\mathcal{M}$; and a set $F(x_2)$ of indexes $i \in V(J)$ such that there is $\Phi(x_2y_1, ijk)$ in $\mathcal{M}$. Note that, from the way we construct $\mathcal{H}$, we know that $F(x_1), F(x_2) \subseteq Y_x$. In view of the above discussion, the number of times x is the endpoint of a path in the 2-matchings of $\mathcal{Q}$ is the number of indexes $i \in Y_x$ such that $i \notin F(x_1) \cap F(x_2)$. Therefore, this number of indexes i such that x is the endpoint of a path in Q_i is at most

$$\begin{aligned}
|Y_x \setminus F(x_1)| + |Y_x \setminus F(x_2)| &\leq 3\alpha n + 2n^{2/3} - 3(|Z_{x_1} \cap \mathcal{M}| + |Z_{x_2} \cap \mathcal{M}|) \\
&\leq 3\alpha n + 2n^{2/3} - 3(\deg_G(x) - \varepsilon' n/2) \\
&\leq 3\alpha n + 2n^{2/3} - 3(\alpha n - n^{1-\rho} - \varepsilon' n/2) \\
&\leq \varepsilon' n
\end{aligned}$$

where in the first inequality we use the facts that $|F(x_1)| = 3|Z_{x_1} \cap \mathcal{M}|$ and $|F(x_2)| = 3|Z_{x_2} \cap \mathcal{M}|$, and also that $Y_x \leq 3\alpha n + n^{2/3}$; in the second inequality we use (4.1); the third inequality follows from $\deg_G(x) \geq \alpha n - n^{1-\rho}$; and since n is sufficiently large, the last inequality holds. This verifies (Q3).

To see (Q4), let $e \in E(Q)$ be arbitrary. As explained before, there exists an orientation $(x, y) \in E(D)$ of e and an edge $ijk \in E(J)$ such that $\Phi(x_1 y_2, ijk)$ belongs to $\mathcal{M}$, and $\{a : e \in E(Q_a)\} = \{i, j, k\} \cap U_1$. Since the latter set obviously has at most three elements, (Q4) follows.

Finally, property (Q5) follows from the properties of the chosen test sets. More precisely, we want to prove that $\Delta(G') \le \varepsilon' n$ for $G' := G - E(Q)$. Since for any $x \in V(G)$ we have $\deg_{G'}(x) = \deg_G(x) - (|Z_{x_1} \cap \mathcal{M}| + |Z_{x_2} \cap \mathcal{M}|)$, it is enough to prove that $|Z_{x_1} \cap \mathcal{M}| + |Z_{x_2} \cap \mathcal{M}| \ge \deg_G(x) - \varepsilon' n$. For that, by using (ii) and (iii) and the facts that $|E(\mathcal{H})| \le |V(\mathcal{H})|\Delta_1(\mathcal{H})/8$ and $|Z_{x_1}| + |Z_{x_2}| \ge \delta_1(\mathcal{H})(|N_B(x_1)| + |N_B(x_2)|) \ge d(1 - d^{-\rho/3}) \deg_G(x)$, we have the following for any $x \in V(G)$:

$$
\begin{aligned}
|Z_{x_1} \cap \mathcal{M}| + |Z_{x_2} \cap \mathcal{M}| &\ge \frac{(1 - d^{-(\rho/3)^3})^2(|Z_{x_1}| + |Z_{x_2}|)|V(\mathcal{H})|/8}{|V(\mathcal{H})|\Delta_1(\mathcal{H})/8} \\
&\ge (1 - d^{-(\rho/3)^3})^2(1 - d^{-\rho/3}) \deg_G(x) \\
&\ge \deg_G(x) - \varepsilon' n/2 \qquad (4.1)
\end{aligned}
$$

where inequality (4.1) holds for sufficiently large n because $\deg_G(x) = \Theta(n)$ and $d = \Theta(n^2)$. Then, we verified that (Q5) holds. This finishes the proof of the lemma. □

5 | Breaking Cycles and Connecting Paths

For a real number $\varepsilon \ge 0$, a collection $\mathcal{P}$ of paths in G is an *ε-almost separating path system* if there exists a set $E' \subseteq E(G)$ such that $\mathcal{P}$ separates every edge in E' from all other edges in G and $\Delta(G - E') \le \varepsilon n$. Note that such $\mathcal{P}$ strongly separates E'.

A 2-matching that has no cycle is called *acyclic*. A collection $\mathcal{Q}$ of 2-matchings is *acyclic* if each 2-matching in $\mathcal{Q}$ is acyclic. The next lemma shows that a collection of 2-matchings as in the output of Theorem 4.2 (that is, a separator, as in Theorem 4.1) can be converted into an acyclic 2-matching with only a very small loss in its properties.

Lemma 5.1. *Let $1/n \ll \delta, L, \beta, \varepsilon$. If G is an n-vertex graph and there exists a $(\delta, L, \beta, \varepsilon)$-separator for G, then there also exists an acyclic $(\delta, L, \beta, 5\varepsilon)$-separator for G.*

Proof. Let $\mathcal{Q}$ be a $(\delta, L, \beta, \varepsilon)$-separator for G. Our goal is to prove that there is a set of edges with at most $4\varepsilon n$ edges incident to each vertex, obtained by deleting one edge from each cycle in the 2-matchings of $\mathcal{Q}$. We argue that this is enough to conclude the proof. First note that, after removing one edge from each cycle of an ε-compact 2-matching Q, we obtain an acyclic 2ε-compact 2-matching, because the maximum number of cycles in Q is εn. Since Q is (δ, L)-robustly-connected, then it remains so after the removal of such edges. Thus, the collection of 2-matchings obtained after the removal of these edges from $\mathcal{Q}$ satisfies (Q1) with 2ε in the place of ε. Note that such collection also satisfies (Q2) and (Q4). Moreover, if we remove from the 2-matchings at most $4\varepsilon n$ edges incident to each vertex of G, then (Q3) will hold with $5\varepsilon n$ in the place of εn. Moreover, the degree of u in $G - E_i$ will increase by at most $4\varepsilon n$, which implies condition (Q5) with 5ε in the place of εn.

Let $C_1, \ldots, C_T$ be the cycles in 2-matchings of $\mathcal{Q}$ and note that $T \le \varepsilon \beta n^2$. For $1 \le i \le T$, let $\mathcal{X}_i$ be the edges of the cycle C_i, and let X_i be an edge chosen uniformly at random from $\mathcal{X}_i$.

Let S be the edge set $\{X_1, \ldots, X_T\}$ and, for each vertex u, let $f^u(X_1, \ldots, X_T)$ be the degree of u in $G[S]$. Note that, since u is in at most one cycle C_i of a 2-matching of $\mathcal{Q}$ and each cycle has length at least $1/\varepsilon$, the edge X_i was chosen as one of the two edges incident to u with probability at most 2ε. Then, because the number of 2-matchings is at most βn, we have that $\mathbf{E}[f^u(X_1, \ldots, X_T)] \le 2\varepsilon \beta n$.

Let $(x_1, \ldots, x_T)$ and $(x_1', \ldots, x_T')$ be in $\mathcal{X}_1 \times \cdots \times \mathcal{X}_T$, differing in exactly one coordinate, that is, $x_j = x_j'$ for every $j \in \{1, \ldots, T\}$ with $j \ne i$. Note that f is such that $|f^u(x_1, \ldots, x_T) - f^u(x_1', \ldots, x_T')| \le 1$. In fact, $f^u(x_1, \ldots, x_T) = f^u(x_1', \ldots, x_T')$ if u is not in C_i. So we can set $c_i = 1$ if u is in C_i and $c_i = 0$ otherwise. As u is in at most βn of the T cycles, we have that $\sum_j c_j^2 \le \beta n$. By using McDiarmid's inequality (Theorem 3.3), we obtain

$$
\Pr\left[f^u(X_1, \ldots, X_T) \ge 4\varepsilon \beta n\right] \le \exp\left(-\frac{8\varepsilon^2 \beta^2 n^2}{\beta n}\right) \le \exp\left(-8\varepsilon^2 \beta n\right)
$$

Thus, by the union bound, the probability that the maximum degree in $G[S]$ is less than $4\varepsilon \beta n$ is at least $1 - n \cdot \exp\left(-8\varepsilon^2 \beta n\right)$, which, for large enough n, is positive. This means there is a choice of edges whose removal from the cycles in the 2-matchings of $\mathcal{Q}$ makes $\mathcal{Q}$ acyclic and satisfying (Q1)–(Q5), with 5ε in the place of ε. □

The next lemma is the main result of this section, and will be used in Section 7 to prove our main result. In this lemma, we will construct an ε-almost separating path system from a separator, which is a collection of 2-matchings. By Lemma 5.1, we can assume the given collection of 2-matchings is acyclic, so we need to transform each path collection into a path by merging its paths one by one into a long path. This merging process can be done by repeated applications of the (δ, L)-robustly-connected property, but each such application adds new edges to the current path collection and thus might affect the original separating property in certain places. This means we need to track carefully the effects of adding a path for every time we modify the current path collection.

Lemma 5.2. *For each ε, δ, and L, there exist ε' and n_0 such that the following holds for every n-vertex graph G with $n \geq n_0$ and every $\beta \in (0, 1)$. If $\mathcal{Q}$ is a $(\delta, L, \beta, \varepsilon')$-separator for G, then there exists an ε-almost separating path system with βn paths.*

Proof. Let $t = \beta n$. We can apply Lemma 5.1 to transform $\mathcal{Q}$ into an acyclic collection of 2-matchings, adjusting the value of ε' accordingly. Let $\mathcal{Q} = \{Q_1, \ldots, Q_t\}$.

We will describe a sequence $\mathcal{C}_0, \ldots, \mathcal{C}_t$ of collections of acyclic 2-matchings in G and sets $E_0, \ldots, E_t$ of edges of G, the idea being that $\mathcal{C}_i$ strongly separates E_i, and that each $\mathcal{C}_i$ will be obtained from $\mathcal{C}_{i-1}$ by replacing Q_i with a path P_i. Then $\mathcal{C}_t$ will be the desired path system.

For each vertex u and $0 \leq i \leq t$, let $d_i(u)$ be the total number of paths in the 2-matchings in $Q_{i+1}, \ldots, Q_t$ that have u as an endpoint. We will make sure the following invariants on $\mathcal{C}_i$ and E_i hold for each $0 \leq i \leq t$:

- I1. each $\mathcal{C}_i$ separates every edge in E_i from all other edges of G;
- I2. edges in more than three of the 2-matchings in $\mathcal{C}_i$ are in $E(\mathcal{C}_i) \setminus E(\mathcal{C}_0)$; and
- I3. the degree of each vertex u in $G - E_i$ is at most εn if $d_i(u) = 0$ and at most $\sqrt{\varepsilon'}n - 2d_i(u)$ if $d_i(u) > 0$.

Let $E_0 = E(\mathcal{Q})$ and $\mathcal{C}_0 = \mathcal{Q}$. Note that $\mathcal{C}_0$ and E_0 satisfy the three invariants. We will define $\mathcal{C}_i = (\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{P_i\}$ for $i = 1, \ldots, t$, where P_i is a path that contains all paths in Q_i. Therefore, if invariants (I1) and (I3) hold for $i = t$ and $\varepsilon' \leq \varepsilon^2$, then $\mathcal{C}_t$ will be an ε-almost separating path system with t paths, and the proof of the lemma will be complete, as $t = \beta n$.

Suppose $i \geq 1$. To describe how we build P_i from Q_i, we need some definitions. Let f be an edge of G not in Q_i such that $Q_i + f$ is a 2-matching. Let E^f be f plus the set of edges of Q_i in E_{i-1} that are not separated from f by $(\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{Q_i + f\}$.

Claim 5.3. If f is in at most three of the 2-matchings in $\mathcal{C}_{i-1}$, then $|E^f| \leq 4$. □

Proof. Suppose there are three edges a, b, and c in $E(Q_i) \cap E_{i-1}$ that are not separated from f by $(\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{Q_i + f\}$. By invariant (I1), there are 2-matchings $Q_{a\bar{b}}$, $Q_{b\bar{c}}$, and $Q_{c\bar{a}}$ in $\mathcal{C}_{i-1}$ such that $a \in E(Q_{a\bar{b}})$ but $b \notin E(Q_{a\bar{b}})$, $b \in E(Q_{b\bar{c}})$ but $c \notin E(Q_{b\bar{c}})$, and $c \in E(Q_{\bar{a}c})$ but $a \notin E(Q_{\bar{a}c})$. Clearly, these three 2-matchings are distinct and are not Q_i, because a, b, and c are in Q_i. So they are in $(\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{Q_i + f\}$ and they must contain f because a, b, and c are not separated from f in $(\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{Q_i + f\}$. By the hypothesis of the claim, these are the only 2-matchings in $\mathcal{C}_{i-1}$ containing f. Hence, repeating the argument for a, b, c in the inverse order, we deduce that either $a \in Q_{b\bar{c}}$, $b \in Q_{\bar{a}c}$, and $c \in Q_{a\bar{b}}$, or $a \notin Q_{b\bar{c}}$, $b \notin Q_{\bar{a}c}$, and $c \notin Q_{a\bar{b}}$.

Now, suppose there is a fourth edge d in $E(Q_i) \cap E_{i-1}$ not separated from f by the collection $(\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{Q_i + f\}$. Consider the former of the two cases above and, for clarity, rename the three 2-matchings to $Q_{ab\bar{c}}$, $Q_{\bar{a}bc}$, and $Q_{a\bar{b}c}$. Then d must be in $Q_{\bar{a}bc}$, to be separated from a, and d must be in $Q_{a\bar{b}c}$, to be separated from b. But now there is no way to separate c from d, a contradiction. The other case is analogous. Indeed, for clarity, rename the three 2-matchings to $Q_{\bar{a}b\bar{c}}$, $Q_{\bar{a}\bar{b}c}$, and $Q_{a\bar{b}\bar{c}}$. Then d must not be in $Q_{\bar{a}b\bar{c}}$ so that b is separated from d, and d must not be in $Q_{a\bar{b}\bar{c}}$ so that a is separated from d. But now there is no way to separate d from c, a contradiction. □

A vertex u is *tight* if its degree in $G - E_{i-1}$ is more than $\varepsilon n - 2$ if $d_{i-1}(u) = 0$, or more than $\sqrt{\varepsilon'}n - 2d_{i-1}(u) - 2$ if $d_{i-1}(u) > 0$. An edge f is *available for* P_i if $f \notin E(\mathcal{C}_{i-1}) \setminus E(\mathcal{C}_0)$ and the extremes of the edges in E^f are not tight.

To transform Q_i into P_i, we will proceed as follows. Start with P_i' being one of the paths in Q_i and let $Q_i' = Q_i \setminus \{P_i'\}$. While Q_i' is nonempty, let P be one of the paths in Q_i'. Call y one of the ends of P and x one of the ends of P_i'. An (x, y)-path in G is *good* if it has length at most L, all of its edges are available and its inner vertices are not in $V(P_i) \cup V(Q_i')$. If a good (x, y)-path exists, we extend P_i' by gluing P_i' and P; we remove P from Q_i', and repeat this process until Q_i' is empty. When Q_i' is empty, we let P_i be P_i'. Recall that $\mathcal{C}_i = (\mathcal{C}_{i-1} \setminus \{Q_i\}) \cup \{P_i\}$, and we let E_i be E_{i-1} minus all edges contained in more than three 2-matchings of $\mathcal{C}_i$ and all edges not separated by $\mathcal{C}_i$ from some other edge of G.

This process is well-defined if the required (x, y)-good path exists at every point in the construction. We will show that, indeed, assuming that the invariants hold, there is always a good (x, y)-path to be chosen in the gluing process above. Then, to complete the proof, we will prove that the invariants hold even after Q_i is modified by the choice of any good path.

First, note that the number of vertices in P_i' not in Q_i is less than $L\varepsilon' n$. Indeed, each connecting path has at most L inner vertices and Q_i is ε'-compact, hence Q_i has no more than $\varepsilon' n$ paths. Thus, we use less than $\varepsilon' n$ connecting paths to get to P_i. If $\varepsilon' < \delta/(4L)$, then the number of vertices in P_i' not in Q_i is less than $\delta n/4$.

Second, let us consider the tight vertices. We start by arguing that x is not tight. This happens because $d_i(x) = d_{i-1}(x) - 1$ and, by invariant (I3), the degree of x in $G - E_{i-1}$ is at most $\sqrt{\varepsilon'}n - 2d_{i-1}(x) = \sqrt{\varepsilon'}n - 2d_i(x) - 2$. For the same reasons, y is not tight. Now, note that $E_i \setminus E_{i-1} \subseteq \bigcup\{E^f : f \in E(P_i) \setminus E(Q_i)\}$. Hence, $|E_i \setminus E_{i-1}| \leq 4L\varepsilon' n$ by Claim 5.3 and because Q_i consists of at most $\varepsilon' n$ paths. This, $\Delta(G - E(Q)) \leq \varepsilon' n$, and $d_i(G) \leq \varepsilon' n$ imply that the maximum number of tight vertices is at most $(\varepsilon' n + 4L\varepsilon'\beta n)/(\sqrt{\varepsilon'} - 2\varepsilon') = \varepsilon'(1 + 4L\beta)n/(\sqrt{\varepsilon'} - 2\varepsilon')$. As long as $2\varepsilon' < \sqrt{\varepsilon'}/2$, that is, $\varepsilon' < 1/16$, we have that this number is less than $2\sqrt{\varepsilon'}(1 + 4L\beta)n$. If additionally $\varepsilon' < (\delta/(8(1 + 4L\beta)))^2$, we have that the number of tight vertices is less than $2\sqrt{\varepsilon'}(1 + 4L\beta)n < \delta n/4$.

Third, $|E(C_{i-1}) \setminus E(C_0)| < 4L\varepsilon' n(i - 1) < 4L\varepsilon'\beta n^2$ because $i \leq \beta n$. Hence, by invariant (I2), at most $4L\varepsilon'\beta n^2$ edges are used more than three times by C_{i-1}. Let $e \in E(C_{i-1}) \setminus E(C_0)$. Because Q_i is (δ, L)-robustly-connected, there exist $\ell \leq L$ and δn^ℓ (x, y)-paths in G, each with ℓ internal vertices, all in $V(G) \setminus V(Q_i)$. If e is not incident to x or y, then the number of (x, y)-paths in G with ℓ internal vertices and containing e is at most $n^{\ell-2}$. Hence, the number of (x, y)-paths in G with ℓ internal vertices, containing an edge in $E(C_{i-1}) \setminus E(C_0)$ not incident to x or y, is less than $4L\varepsilon'\beta n^\ell$. If e is incident to x or y, then the number of (x, y)-paths in G with ℓ internal vertices and containing e is at most $n^{\ell-1}$. But, there are less than $\sqrt{\varepsilon'}n$ edges incident to x and less than $\sqrt{\varepsilon'}n$ edges incident to y contained in more than three 2-matchings in C_{i-1}, by invariant (I3). Thus, the number of (x, y)-paths in G of length ℓ containing an edge in $E(C_{i-1}) \setminus E(C_0)$ incident to x or y is less than $2\sqrt{\varepsilon'}n^\ell$. We can choose ε' small enough so that $4L\varepsilon'\beta + 2\sqrt{\varepsilon'} < \delta/4$, and thus at most $\delta n^\ell/4$ (x, y)-paths of length ℓ contain some edge of $E(C_{i-1}) \setminus E(C_0)$.

Summarizing, we have concluded that, for ε' small enough, the number of vertices in P_i' not in Q_i is less than $\delta n/4 \leq \delta n^\ell/4$, the number of tight vertices is also less than $\delta n/4 \leq \delta n^\ell/4$, and the number of (x, y)-paths containing some edge in $E(C_{i-1}) \setminus E(C_0)$ is less than $\delta n^\ell/4$. This means that at least $\delta n^\ell/4$ of the δn^ℓ (x, y)-paths in G, each with ℓ internal vertices, all in $V(G) \setminus V(Q_i)$, are good. As long as n_0 is such that $\delta n_0^\ell/4 \geq \delta n_0/4 \geq 1$, there is a good (x, y)-path.

Now, let us verify the invariants. By the definition of E_i, invariant (I1) holds for i because C_i separates every edge in E_i from all other edges of G. Invariant (I2) holds because edges in more than three 2-matchings in C_i lie in used connecting paths, that is, lie in $E(P_j) \setminus E(Q_j)$ for some j with $1 \leq j \leq i$. For invariant (I3), observe that $E_i \setminus E_{i-1} \subseteq E(P_i)$, so the degree of v from $G - E_{i-1}$ to $G - E_i$ decreases only for untight vertices, and by at most two. As the degree of an untight vertex u in $G - E_{i-1}$ is at most $\varepsilon n - 2$ if $d_{i-1}(u) = 0$ and at most $\sqrt{\varepsilon'}n - 2d_{i-1}(u) - 2$ if $d_{i-1}(u) > 0$, every vertex u in $G - E_i$ has degree at most εn if $d_i(u) = 0$ and at most $\sqrt{\varepsilon'}n - 2d_i(u)$ if $d_i(u) > 0$, also because $d_i(u) \leq d_{i-1}(u)$. So invariant (I3) holds. □

6 | Separating the Last Few Edges

In this section, we deal with a subgraph H of G, of small maximum degree, whose edges are not separated by the path family obtained in the previous sections. This is done in Theorem 6.3 but first we need some auxiliary results. The first step of the proof is to find a family of matchings which separates the edges of H.

Lemma 6.1. *Let $\Delta \geq 0$ and let H be an n-vertex graph with $\Delta(H) \leq \Delta$. Then there is a collection of $t \leq 300\sqrt{\Delta n}$ matchings $M_1, \ldots, M_t \subseteq H$ such that*

- *M1. each edge in H belongs to exactly two matchings M_i, M_j; and*
- *M2. for each $1 \leq i < j \leq t$, the matchings M_i, M_j have at most one edge in common.*

We also need the asymmetric version of the Lovász Local Lemma (cf. [18], Theorem 1.1).

Theorem 6.2 (Asymmetric Lovász Local Lemma). *Let $\mathcal{E} = \{A_1, \ldots, A_n\}$ be a collection of events such that each A_i is mutually independent of $\mathcal{E} - (\mathcal{D}_i \cup A_i)$, for some $\mathcal{D}_i \subseteq \mathcal{E}$. Let $0 < x_1, \ldots, x_n < 1$ be real numbers such that, for each $i \in \{1, \ldots, n\}$,*

$$\Pr[A_i] \leq x_i \prod_{A_j \in \mathcal{D}_i} (1 - x_j) \tag{6.1}$$

Then $\Pr\left[\bigcap_{i=1}^n \overline{A_i}\right] \geq \prod_{i=1}^n (1 - x_i) > 0$.

Proof of Theorem 6.1. Let $D = 256\sqrt{\Delta n}$. Let $M = \{1, \ldots, D + 1\}$ and let $M^{(2)}$ consist of all subsets of size two of M. We define a function $\phi : E(H) \to M^{(2)}$ by choosing $\phi(e) \in M^{(2)}$ uniformly at random for each $e \in E(H)$. We will show that, with positive probability,

i. ϕ is injective, and

ii. for each vertex $v \in V(H)$, the sets $\phi(vw)$ for $w \in N(v)$ are pairwise disjoint.

We define a sequence of "bad" events to use Theorem 6.2. For distinct $e, f \in E(H)$, let $\mathcal{A}_{e,f}$ be the event that $\phi(e) = \phi(f)$. For each pair of adjacent edges $e, f \in E(H)$, let $\mathcal{B}_{e,f}$ be the event that $\phi(e) \cap \phi(f) \neq \emptyset$. Thus (i)–(ii) hold if we avoid all $\mathcal{A}_{e,f}$ and $\mathcal{B}_{e,f}$.

Note first that, for each e, f, we have

$$\Pr[\mathcal{A}_{e,f}] = \binom{D+1}{2}^{-1} = \frac{2}{D(D+1)} \leq \frac{2}{D^2}$$
$$\Pr[\mathcal{B}_{e,f}] = (2D-1)\binom{D+1}{2}^{-1} = \frac{2(2D-1)}{D(D+1)} \leq \frac{4}{D}$$

Define $d_A := \Delta n$ and $d_B := 4\Delta$. Note that each event $\mathcal{A}_{e,f}$ or $\mathcal{B}_{e,f}$ is independent of all other events $\mathcal{A}_{e',f'}$ except if $\{e, f\} \cap \{e', f'\} \neq \emptyset$. Given $\{e, f\}$, the number of such intersecting pairs $\{e', f'\}$ is at most d_A. Similarly, each event $\mathcal{A}_{e,f}$ or $\mathcal{B}_{e,f}$ is independent of all but at most d_B events of type $\mathcal{B}_{e',f'}$.

For each event $\mathcal{A}_{e,f}$ define $x_{e,f} := x_A := d_A^{-1}$ and for each event $\mathcal{B}_{e,f}$ define $y_{e,f} := x_B := d_B^{-1}$. We will show that the requirement (6.1) of the Asymmetric Lovász Local Lemma is satisfied with these choices.

Indeed, for an event of type $\mathcal{A}_{e,f}$, we use the fact that $1 - x \geq 2^{-2x}$ for $0 \leq x \leq 1/2$ to show that

$$\begin{aligned} x_A (1 - x_A)^{d_A} (1 - x_B)^{d_B} &\geq x_A 2^{-2x_A d_A} 2^{-2x_B d_B} \\ &= x_A 2^{-4} = \frac{1}{16\Delta n} \geq \frac{2}{D^2} \\ &\geq \Pr[\mathcal{A}_{e,f}] \end{aligned}$$

and, for an event of type $\mathcal{B}_{e,f}$, we have

$$\begin{aligned} x_B (1 - x_A)^{d_A} (1 - x_B)^{d_B} &\geq x_B 2^{-2x_A d_A} 2^{-2x_B d_B} \\ &= x_B 2^{-4} = \frac{1}{64\Delta} \geq \frac{4}{D} \geq \Pr[\mathcal{B}_{e,f}] \end{aligned}$$

Thus, Theorem 6.2 guarantees there is a function ϕ satisfying (i)–(ii). This function defines the matchings: for each $1 \leq i \leq D+1$ we let M_i consist of the edges $e \in E(H)$ such that $i \in \phi(e)$. Then $\phi(e) \in M^{(2)}$ ensures that each edge belongs to exactly two M_i's, condition (i) ensures that each pair of M_i, M_j has at most one edge in common, and condition (ii) ensures that each M_i is a matching. Since $D + 1 \leq 300\sqrt{\Delta n}$, we are done. □

Now, we prove the main result of this section, which finds the required family of paths that separate $E(H)$. The proof proceeds by using the matchings found in the previous lemma and covering those matchings with paths. We note that our task here is substantially easier than in Lemma 5.2 (where we also needed to extend a path family into a single path by adding new paths) because here the connecting paths are found outside the set of edges we are trying to separate.

Lemma 6.3. *Let $\varepsilon, \delta, L > 0$ and let G and H be n-vertex graphs with $H \subseteq G$ such that $\Delta(H) \leq \varepsilon n$ and G is (δ, L)-robustly-connected. Then there exist paths $\{P_i\}_{i=1}^r$, $\{Q_i\}_{i=1}^r$ in G, with $r \leq 600L\delta^{-1}\sqrt{\varepsilon}n$, such that, for each $e \in E(H)$, there exist distinct $1 \leq i < j \leq r$ such that $\{e\} = E(P_i) \cap E(P_j) \cap E(Q_i) \cap E(Q_j)$.*

Proof. Apply Theorem 6.1 to H (with εn in place of Δ), to obtain a collection of $t \leq 300\sqrt{\varepsilon}n$ matchings $M_1, \dots, M_t$ such that each edge in H belongs to exactly two of these matchings; and each two distinct matchings have at most one edge in common.

Separate the edges of each M_i into $r_i \leq 2L\delta^{-1}$ matchings $M_{i,1}, \dots, M_{i,r_i}$ where each $M_{i,j}$ has less than $\delta n/(4L)$ edges. Let $r = \sum_i r_i$ be the total number of matchings obtained after doing this. Since we have t matchings M_i initially, after this process, we have obtained at most $r \leq t2L\delta^{-1} \leq 600L\delta^{-1}\sqrt{\varepsilon}n$ matchings $M_{i,j}$. We rename and enumerate the new matchings to be $M'_1, \dots, M'_r$ from now on.

The next step is to obtain, for each $1 \leq i \leq r$, two paths P_i and Q_i of G with the property that $E(P_i) \cap E(Q_i) = M'_i$. For that, let $x_1y_1, x_2y_2, \dots, x_\ell y_\ell$ be the edges of M'_i. Since G is (δ, L)-robustly-connected, there exist $1 \leq \ell \leq L$ and at least δn^ℓ many internally vertex-disjoint (x_1, x_2)-paths with ℓ inner vertices each. Because $|V(M'_i)| = 2|M'_i| \leq \delta n/(2L) < \delta n$, there exists an (x_1, x_2)-path $P_i^{(1)}$ of length at most L which is internally vertex-disjoint from $V(M'_i)$. Similarly, we can find a (y_1, y_2)-path $Q_i^{(1)}$ of length at most L which is internally disjoint from $V(M'_i) \cup V(P_i^{(1)})$. We proceed in this fashion iteratively, finding for each $1 \leq k < \ell$, in order, some (x_k, x_{k+1})-path $P_i^{(k)}$ and a (y_k, y_{k+1})-path $Q_i^{(k)}$, both of length at most L, and both internally disjoint from $V(M_i')$ and from all previously chosen paths. This can be achieved, because in each step the number of vertices we need to avoid is at most $2|M_i'| + 2L|M_i'| \leq 4L|M_i'| < \delta n$, which implies that there is always one path available to choose. We define P_i as the path which starts with the edge x_1y_1, then traverses the path $Q_i^{(1)}$, then uses y_2x_2, then $P_i^{(2)}$, etc., alternatingly using the paths $P_i^{(k)}$ and $Q_i^{(k)}$, and covering all edges of M_i'. We define Q_i similarly, starting by the edge y_1x_1, but then using the path $P_i^{(1)}$, then x_2y_2, then $Q_i^{(2)}$, and so on. Then P_i, Q_i satisfy that $E(P_i) \cap E(Q_i) = M_i'$, as required.

We define $\mathcal{P} = \{P_1, \dots, P_r\}$ and $\mathcal{Q} = \{Q_1, \dots, Q_r\}$. By construction, each of them has the required number of paths. Now we check that these families satisfy the required property. Let $e \in E(H)$ be arbitrary. By the choice of the matchings, there exist distinct i_1, i_2 such that $\{e\} = M_{i_1} \cap M_{i_2}$. Suppose that i, j are distinct such that $e \in M_i' \subseteq M_{i_1}$ and $e \in M_j' \subseteq M_{i_2}$. It must happen that $\{e\} = M_i' \cap M_j'$. Then, by the choice of P_i, Q_i, P_j, Q_j, we have that $M_i' = E(P_i) \cap E(Q_i)$ and $M_j' = E(P_j) \cap E(Q_j)$, and therefore $E(P_i) \cap E(Q_i) \cap E(P_j) \cap E(Q_j) = M_i' \cap M_j' = \{e\}$, as required. □

7 | Proof of the Main Result

Now, we have the tools to prove our main result, from which Theorems 1.1 and 1.2 immediately follow (in combination with the lower bounds from Theorems 2.1 and 2.3).

Theorem 7.1. *Let $\alpha, \rho, \varepsilon, \delta \in (0,1)$ and $L > 0$. Let n be sufficiently large, and let G be an n-vertex $(\alpha n \pm n^{1-\rho})$-regular graph which is (δ, L)-robustly-connected. Then $\mathrm{ssp}(G) \leq (\sqrt{3\alpha + 1} - 1 + \varepsilon)n$.*

Proof. Let $\varepsilon_2 := 1 - 1/(1 + \varepsilon/2)$ and $\delta' = \varepsilon_2^{\ell}\delta/2$. Choose ε' and n_0 such that Lemma 5.2 holds with $(\varepsilon\delta/(2400L))^2$, L and δ' playing the roles of ε, L and δ, respectively. From now on, we assume $n \geq n_0$ and let $\beta := \sqrt{3\alpha + 1} - 1$. Apply Theorem 4.2 to G with ε_2 and ε' playing the roles of ε and ε', respectively. By doing this, we obtain a family $\mathcal{Q}$ of 2-matchings which is a $(\delta', L, (1 + \varepsilon/2)\beta, \varepsilon')$-separator. Thus $\mathcal{Q}$ consists of $t := (1 + \varepsilon/2)\beta n \leq (\sqrt{3\alpha + 1} - 1 + \varepsilon/2)n$ many 2-matchings $Q_1, \dots, Q_t$, satisfying (Q1)–(Q5) (with δ', $(1 + \varepsilon/2)\beta$ and ε' in place of δ, β and ε). Next, we apply Lemma 5.2 to G and $\mathcal{Q}$. By the choice of ε' and n_0, we obtain an $(\varepsilon\delta/(2400L))^2$-almost separating path system $\mathcal{P}$ in G of size t.

Let $E' \subseteq E(G)$ be the subset of edges which are strongly separated by $\mathcal{P}$ from every other edge. Since $\mathcal{P}$ is $(\varepsilon\delta/(2400L))^2$-almost separating, the subgraph $J := G - E'$ satisfies $\Delta(J) \leq (\varepsilon\delta/(2400L))^2 n$. By assumption, G is (δ, L)-robustly-connected, which allows us to apply Theorem 6.3 with J and $(\varepsilon\delta/(2400L))^2$ playing the roles of H and ε, respectively. By doing so, we obtain two families $\mathcal{R}_1, \mathcal{R}_2$ of at most $\varepsilon n/4$ paths each, such that, for each $e \in E(J)$, there exist two paths $P_i, P_j \in \mathcal{R}_1$ and $Q_i, Q_j \in \mathcal{R}_2$ such that $\{e\} = E(P_i) \cap E(P_j) \cap E(Q_i) \cap E(Q_j)$.

We let $\mathcal{P}' := \mathcal{P} \cup \mathcal{R}_1 \cup \mathcal{R}_2$. Note that $\mathcal{P}'$ has at most $t + \varepsilon n/2 \leq (\sqrt{3\alpha + 1} - 1 + \varepsilon)n$ many paths. We claim that $\mathcal{P}'$ is a strong-separating path system for G. Indeed, let e, f be distinct edges in $E(G)$; we need to show that there exists a path in $\mathcal{P}'$ which contains e and not f. If $e \in E'$, then such a path is contained in $\mathcal{P}$, so we can assume that $e \in E(J)$. There exist four paths $P_i, P_j, Q_i, Q_j \in \mathcal{P}'$ such that $\{e\} = E(P_i) \cap E(P_j) \cap E(Q_i) \cap E(Q_j)$, which in particular implies that one of these paths does not contain f. □

8 | Corollaries

Now, we apply Theorem 7.1 to bound $\mathrm{ssp}(G)$ for graphs G belonging to certain families of graphs. In all cases, we just need to check that the corresponding graphs are (δ, L)-robustly-connected for suitable parameters.

We begin by considering complete balanced bipartite graphs. Previously, Wickes [19], Chap 9 studied upper and lower bounds for $\mathrm{ssp}(K_{n/2}, K_{n/2})$, and obtained the lower bound $\mathrm{ssp}(K_{n/2}, K_{n/2}) \geq (\sqrt{5/2} - 1)n - 1/2$, which coincides with the main term from our lower bound from Theorem 2.3. We can obtain a corresponding upper bound, which is then asymptotically tight.

Corollary 8.1. *For each $\varepsilon > 0$ and sufficiently large n, $\mathrm{ssp}(K_{n/2,n/2}) \leq (\sqrt{5/2} - 1 + \varepsilon)n$.*

Proof. Let $\varepsilon > 0$ be arbitrary and n sufficiently large in terms of ε. The graph $K_{n/2,n/2}$ is $n/2$-regular, so it is αn-regular with $\alpha = 1/2$. Pairs of vertices x, y in the same part of the bipartition have $n/2$ neighbors in common; and pairs of vertices x, y in different parts of the bipartition have $((n/2) - 1)^2 \geq n^2/5$ many (x, y)-paths with two inner vertices each. Hence, $K_{n/2,n/2}$ is $(1/5, 2)$-robustly-connected. By applying Theorem 7.1 with $\alpha = 1/2$, ε, ρ and $\delta = 1/5$ we obtain that $\mathrm{ssp}(K_n) \leq (\sqrt{5/2} - 1 + \varepsilon)n$. □

Let us now describe a well-known family of graphs which satisfies the connectivity assumptions of Theorem 7.1. Given $0 < \nu \leq \tau \leq 1$, a graph G on n vertices, and a set $S \subseteq V(G)$, the *ν-robust neighborhood of S* is the set $\mathrm{RN}_{\nu,G}(S) \subseteq V(G)$ of all vertices with at least νn neighbors in S. We say that G is a *robust (ν, τ)-expander* if, for every $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$, we have $|\mathrm{RN}_{\nu,G}(S)| \geq |S| + \nu n$.

Many families of graphs are robust (ν, τ)-expanders for suitable values of ν, τ, including large graphs with $\delta(G) \geq dn$ for fixed $d > 1/2$, dense random graphs, dense regular quasirandom graphs [20], Lemma 5.8, etc.

Corollary 8.2. *For each $\varepsilon, \alpha, \tau, \nu, \rho > 0$ with $\alpha \geq \tau + \nu$, there exists n_0 such that the following holds for each $n \geq n_0$. Let G be an n-vertex $(\alpha n \pm n^{1-\rho})$-regular robust (ν, τ)-expander. Then $\mathrm{ssp}(G) \leq (\sqrt{3\alpha + 1} - 1 + \varepsilon)n$.*

Proof. By Theorem 7.1, it is enough to prove that G is (δ, L)-robustly-connected with δ, L depending on ν only. We will prove this holds with $L := \lceil \nu^{-1} \rceil$ and $\delta := (\nu/4)^L 4^{-L^2}$.

Let x be any vertex, and let $N(x)$ be its neighborhood. We define $R_0 = \emptyset$ and for each $i \geq 0$ we let $R_{i+1} = R_i \cup (\mathrm{RN}_{\nu,G}(N(x) \cup R_i) \setminus N(x))$ if $|N(x) \cup R_i| \leq (1-\tau)n$; or $R_{i+1} = V(G)$ otherwise. By definition, $R_0 \subseteq R_1 \subseteq R_2 \subseteq \cdots$.

Since G is a robust (ν, τ)-expander and $|N(x)| \geq \alpha n \geq \tau n$, it can be quickly checked that, for each $i \geq 0$ such that $|N(x) \cup R_i| \leq (1-\tau)n$, the bound $|R_{i+1} \setminus R_i| \geq \nu n$ holds. In particular, this implies that $R_L = V(G)$. Indeed, suppose otherwise. Then $R_L \neq V(G)$, therefore $|N(x) \cup R_i| \leq (1-\tau)n$ for all $0 \leq i < L$, which implies that $|R_i \setminus R_{i-1}| \geq \nu n$ holds for each $1 \leq i \leq L$. But then, since $L \geq \nu^{-1}$, we have

$$n > |R_L| = |R_L \setminus R_{L-1}| + \cdots + |R_2 \setminus R_1| + |R_1| \geq L\nu n \geq n$$

a contradiction.

Given $j \geq 1$, we let $T_j \subseteq V(G)$ be the set of vertices v for which there are at least $(\nu n/4)^j 4^{-j^2}$ many (x, v)-paths in G with j inner vertices each. We claim that, for each $0 \leq i \leq L$, it holds that $R_i \subseteq T_1 \cup \cdots \cup T_i$. Before proving the claim we note that this is enough to conclude: as discussed before we have that $V(G) = R_L \subseteq T_1 \cup \cdots \cup T_L$, so for each vertex $y \in V(G)$ there would exist $1 \leq \ell \leq L$ such that $y \in T_\ell$. This implies that there exist at least $(\nu n/4)^\ell 4^{-\ell^2} \geq \delta n^\ell$ many (x, y)-paths with ℓ inner vertices each, as required.

Now, we prove the claim by induction on i, where the base case $i = 0$ holds vacuously. Assuming the claim for some $i < L$, we prove it for $i+1$. Let $y \in R_{i+1}$ be arbitrary, it is enough to check that $y \in T_1 \cup \cdots \cup T_{i+1}$. By the inductive hypothesis, we can assume that $y \in R_{i+1} \setminus R_i$. Note that y has at least νn neighbors in $N(x) \cup R_i$. Indeed, if $|N(x) \cup R_i| \leq (1-\tau)n$ then $y \in R_{i+1} \setminus R_i \subseteq \mathrm{RN}_{\nu,G}(N(x) \cup R_i)$ so indeed y must have at least νn neighbors in $N(x) \cup R_i$. Otherwise, if $|N(x) \cup R_i| > (1-\tau)n$ then, since y has at least $\alpha n \geq (\tau + \nu)n$ neighbors, at least νn of them must be in $N(x) \cup R_i$.

We are done if y has at least $\nu n/2$ neighbors in $N(x)$ because that immediately implies that $y \in T_1$. We assume from now on that $|N(y) \cap N(x)| < \nu n/2$ and therefore $|N(y) \cap R_i| \geq \nu n/2$. By the induction hypothesis, $R_i \subseteq T_1 \cup \cdots \cup T_i$. Observe that there must exist $1 \leq r \leq i$ such that $|N(x) \cap T_r| \geq \nu n/2^{r+1}$, as otherwise we would have $|N(x) \cap R_i| < (\nu n/2) \sum_{r \geq 1} 2^{-r} \leq \nu n/2$, a contradiction. Fix such an r from now on, and we will conclude by showing that $y \in T_{r+1}$.

Indeed, for each $z \in N(y) \cap T_r$ there is a family $\mathcal{P}_z$ of at least $(\nu n/4)^r 4^{-r^2}$ many (x, z)-paths with r inner vertices each. We wish to extend the paths in $\mathcal{P}_z$ by including y to obtain (x, y)-paths with $r+1$ inner vertices each. This can only fail for some $P \in \mathcal{P}_z$ if $y \in V(P)$, but that can happen only for at most rn^{r-1} paths. Since $|\mathcal{P}_z| \geq (\nu n/4)^r 4^{-r^2}$, using that $r \leq L$ and $1/n \ll \nu$ we can deduce that $|\mathcal{P}_z|/2 \geq rn^{r-1}$. This allows us to conclude that there are at least $|\mathcal{P}_z|/2$ many (x, y)-paths with $r+1$ inner vertices which end with zy. By counting the paths for each choice of $z \in N(y) \cap T_r$, the number of desired (x, y)-paths is at least

$$\begin{aligned}
\sum_{z \in N(y) \cap T_r} \frac{|\mathcal{P}_z|}{2} &\geq \frac{|N(y) \cap T_r|}{2} \left(\frac{\nu n}{4} \right)^r 4^{-r^2} \\
&\geq \frac{\nu n}{4 \cdot 2^r} \left(\frac{\nu n}{4} \right)^r 4^{-r^2} \\
&\geq \left(\frac{\nu n}{4} \right)^{r+1} 4^{-(r+1)^2}
\end{aligned}$$

so $y \in T_{r+1}$, as claimed. This finishes the proof. □

9 | Conclusion

9.1 | Separating All n-Vertex Graphs

To determine the maximum of $\mathrm{wsp}(G)$ and $\mathrm{ssp}(G)$ over all n-vertex graphs G remains an interesting problem. Falgas–Ravry, et al. (see [6], Conjecture 1.2 and the remarks afterwards) said that "it is not inconceivable" that $\mathrm{wsp}(G) \leq (1 + o(1))n$ holds for all n-vertex graphs G. We have shown that $\mathrm{ssp}(G) \leq (1 + o(1))n$ holds for dense, regular, sufficiently connected n-vertex graphs. Even the following could be true[2]:

Question 9.1. Does $\mathrm{ssp}(G) \leq (1+o(1))n$ hold for *all* connected n-vertex graphs G?

We need to consider connected graphs for this question, because we have $\mathrm{ssp}(K_4) = 5$, and so the graph G consisting of $n/4$ vertex-disjoint copies of K_4 satisfies $\mathrm{ssp}(G) = 5n/4$.

9.2 | Separating Nonregular or Nonconnected Graphs

It would also be interesting to estimate $\mathrm{wsp}(G)$ and $\mathrm{ssp}(G)$ for graphs not covered by our main result. Complete bipartite graphs $K_{a,b}$ with $a < b$ are an interesting open case. It is also of interest to weaken the conditions in our main result. For instance, can the connectivity conditions in Theorem 7.1 be weakened? Does $\Omega(n)$-vertex-connectivity suffice?

Another way to weaken the connectivity conditions in Theorem 7.1 could be as follows: by a result of Kühn et al. [21], each near-regular and dense graph can be vertex-partitioned into parts which are robust expanders or "bipartite robust expanders" (see [21], Section 5), so one could try to apply our result separately in each part of the partition and then deal with the remaining edges. As we have already seen (by the example shown after Theorem 1.2), the connectivity condition cannot be removed completely. The chief reason behind it is that the lower bound from Theorem 2.3 becomes close to tight whenever most of the paths are close to being Hamiltonian (as can be seen from inspecting that proof). This means we would need a method to join the paths together in a coherent way to make sure they become close to Hamiltonian and still separate most the edges; to make such a strategy work would require some connectivity condition and new ideas.

9.3 | Exact Results, and Orthogonal Cover Decompositions

As mentioned before, we have $\mathrm{ssp}(K_n) = n$ if and only if an ODC with Hamiltonian paths exists for K_n. Gronau, Müllin, and Rosa [22] conjectured that an ODC by H in K_n can be found whenever H is any n-vertex tree which is not a path with three edges. If true, this would imply that $\mathrm{ssp}(K_n) = n$ holds for every $n \neq 4$. An approximate version of this conjecture (obtained as a corollary of general results about "rainbow trees") was obtained by Montgomery, Pokrovskiy and Sudakov [23], Theo 1.7 whenever n is a large power of two.

It would be interesting to see if Theorem 1.1 could be deduced from the partial known results on ODCs, but we note that we do not see an easy reduction here. For our approach to work it is crucial that the "leftover" graph has bounded maximum degree to be able to separate it with few extra edges (as we do in Section 6), and we did not see a way to obtain this from the known results.

Acknowledgments

This work began at the ChiPaGra workshop 2023 in Valparaíso, Chile, funded by FAPESP-ANID Investigación Conjunta grant 2019/13364-7. We thank the organizers and the attendees for a great atmosphere. The third author thanks Matías Pavez-Signé for useful suggestions.

Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Endnotes

[1] This corresponds to the $j = 1$ and $\varepsilon = \rho$ case of the definition of (d, ε, C)-*trackable test systems* of [14], Section 3. The original definition requires more properties but reduces to the definition we have given when $j = 1$. In particular, C does not play a role anymore, so we opted for removing it from the notation.

[2] In [7], Theorem 10 it is stated that for each $\varepsilon \in (0, 1/2)$ there exists some n and an n-vertex graph G such that $\mathrm{ssp}(G) \geq 2(1 - 2\varepsilon)n$, but unfortunately the proof has a flaw. The error in the proof appears in [7], Remark 10, because the length of the longest path in $K_{\varepsilon n,(1-\varepsilon)n}$ is $2\varepsilon n$ and not $\varepsilon n + 1$.

References

1. A. Rényi, "On Random Generating Elements of a Finite Boolean Algebra," *Acta Scientiarum Mathematicarum* 22 (1961): 75–81, http://acta.bibl.u-szeged.hu/13920/1/math_022_fasc_001_002_075-081.pdf.

2. N. J. A. Harvey, M. Patrascu, Y. Wen, S. Yekhanin, and V. W. S. Chan, "Non-Adaptive Fault Diagnosis for all-Optical Networks via Combinatorial Group Testing on Graphs," in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communi- Cations* (IEEE, 2007), 697–705, https://doi.org/10.1109/INFCOM.2007.87.

3. I. Honkala, M. G. Karpovsky, and S. Litsyn, "Cycles Identifying Vertices and Edges in Binary Hypercubes and 2-Dimensional Tori," *Discrete Applied Mathematics* 129, no. 2–3 (2003): 409–419, https://doi.org/10.1016/S0166-218X(02)00579-6.

4. J. Tapolcai, L. Ronyai, and P.-H. Ho, "Link Fault Localization Using bi-Directional M- Trails in all-Optical Mesh Networks," *IEEE Transactions on Communications* 61, no. 1 (2013): 291–300, https://doi.org/10.1109/TCOMM.2012.091812.110582.

5. L. Zakrevski and M. Karpovsky, "Fault-Tolerant Message Routing for Multiprocessors," in *Lecture Notes in Computer Science, International Parallel Processing Symposium 1998*, vol. 1388 (Springer, 1998), 714–730, https://doi.org/10.1007/3-540-64359-1_737.

6. V. Falgas-Ravry, T. Kittipassorn, D. Korndi, S. Letzter, and B. P. Narayanan, "Separating Path Systems," *Journal of Combinatorics* 5, no. 3 (2014): 335–354, https://doi.org/10.4310/JOC.2014.v5.n3.a4.

7. J. Balogh, B. Csaba, R. R. Martin, and A. Pluhr, "On the Path Separation Number of Graphs," *Discrete Applied Mathematics* 213 (2016): 26–33, https://doi.org/10.1016/j.dam.2016.05.022.

8. S. Letzter, "Separating Path Systems of Almost Linear Size," *Transactions of the American Mathematical Society* 377, no. 8 (2024): 5583–5615, https://doi.org/10.1090/tran/9187.

9. M. Bonamy, F. Botler, F. Dross, T. Naia, and J. Skokan, "Separating the Edges of a Graph by a Linear Number of Paths," *Advances in Combinatorics* 6 (2023): 7pp, https://doi.org/10.19086/aic.2023.6.

10. B. Wickes, "Separating Path Systems for the Complete Graph," *Discrete Mathematics* 347, no. 3 (2024): 113784, https://doi.org/10.1016/j.disc.2023.113784.

11. G. Kontogeorgiou and M. Stein, "An Exact Upper Bound for the Minimum Size of a Path System That Weakly Separates a Clique," 2024 10.48550/arXiv.2403.08210.Preprint.

12. H.-D. O. F. Gronau, M. Grüttmüller, S. Hartmann, U. Leck, and V. Leck, "On Orthogonal Double Covers of Graphs," *Designs, Codes and Criptography* 27 (2002): 49–91, https://doi.org/10.1023/A:1016546402248.

13. J. D. Horton and G. M. Nonay, "Self-Orthogonal Hamilton Path Decompositions," *Discrete Mathematics* 97, no. 1–3 (1991): 251–264, https://doi.org/10.1016/0012-365X(91)90441-4.

14. S. Glock, F. Joos, J. Kim, M. Kühn, and L. Lichev, "Conflict-Free Hypergraph Matchings," *Journal of the London Mathematical Society* 109.5, no. 2 (2024): e12899, https://doi.org/10.1112/jlms.12899.

15. M. Delcourt and L. Postle, "Finding an Almost Perfect Matching in a Hypergraph Avoiding Forbidden Submatchings," 2022, https://arxiv.org/abs/2204.08981. Preprint.

16. S. Janson, T. Łuczak, and A. Rucinski, "Random Graphs," in *Wiley-Interscience Series in Discrete Mathematics and Optimization* (Wiley-Interscience, 2000), xii+333, https://doi.org/10.1002/9781118032718.

17. C. McDiarmid, "On the Method of Bounded Differences," in *Surveys in Combinatorics*, vol. 141 (Cambridge University Press, 1989), 148–188, https://doi.org/10.1017/CBO9781107359949.008.

18. J. Spencer, "Asymptotic Lower Bounds for Ramsey Functions," *Discrete Mathematics* 20, no. 1 (1977): 69–76, https://doi.org/10.1016/0012-365X(77)90044-9.

19. B. Wickes, *Covering and Separation for Permutations and Graphs*". PhD thesis (School of Mathematical Sciences, 2023), https://qmro.qmul.ac.uk/xmlui/handle/123456789/93230.

20. D. Kühn and D. Osthus, "Hamilton Decompositions of Regular Expanders: Applications," *Journal of Combinatorial Theory Series B* 104 (2014): 1–27, https://doi.org/10.1016/j.jctb.2013.10.006.

21. D. Kühn, A. Lo, D. Osthus, and K. Staden, "The Robust Component Structure of Dense Regular Graphs and Applications," *Proceedings of the London Mathematical Society* 110, no. 1 (2015): 19–56.

22. H.-D. O. F. Gronau, R. C. Mullin, and A. Rosa, "Orthogonal Double Covers of Complete Graphs by Trees," *Graphs and Combinatorics* 13, no. 3 (1997): 251–262, https://doi.org/10.1007/BF03353002.

23. R. Montgomery, A. Pokrovskiy, and B. Sudakov, "Embedding Rainbow Trees With Applications to Graph Labelling and Decomposition," *Journal of the European Mathematical Society* 22, no. 10 (2020): 3101–3132, https://doi.org/10.4171/jems/982.